

A NONNEGATIVITY PRESERVING SCHEME FOR THE RELAXED CAHN–HILLIARD EQUATION WITH SINGLE-WELL POTENTIAL AND DEGENERATE MOBILITY

FEDERICA BUBBA^{1,2,†} AND ALEXANDRE POULAIN^{1,‡,*} 

Abstract. We propose and analyze a finite element approximation of the relaxed Cahn–Hilliard equation [Perthame and Poulain, *Eur. J. Appl. Math.* **32** (2021) 89–112.] with singular single-well potential of Lennard-Jones type and degenerate mobility that is energy stable and nonnegativity preserving. The Cahn–Hilliard model has recently been applied to model evolution and growth for living tissues. Although the choices of degenerate mobility and singular potential are biologically relevant, they induce difficulties regarding the design of a numerical scheme. We propose a finite element scheme, and we show that it preserves the physical bounds of the solutions thanks to an upwind approach adapted to the finite element method. We propose two different time discretizations leading to a non-linear and a linear scheme. Moreover, we show the well-posedness and convergence of solutions of the non-linear numerical scheme. Finally, we validate our scheme by presenting numerical simulations in one and two dimensions.

Mathematics Subject Classification. 35Q92, 65M60, 35K55, 35K65, 35K35.

Received June 30, 2021. Accepted May 16, 2022.

1. INTRODUCTION

Being of fourth order, the Cahn–Hilliard equation does not fit usual softwares for finite elements. To circumvent this difficulty a relaxed version has been proposed in [32] and the presentation of a finite element numerical scheme that preserves the physical properties of the solutions is the purpose of the present work. The relaxed version of the Cahn–Hilliard (RDCH in short) equation reads

$$\begin{cases} \frac{\partial n}{\partial t} = \nabla \cdot (b(n) \nabla (\varphi + \psi'_+(n))) , \\ -\sigma \Delta \varphi + \varphi = -\gamma \Delta n + \psi'_- \left(n - \frac{\sigma}{\gamma} \varphi \right) , \end{cases} \quad t > 0, \quad x \in \Omega, \quad (1.1)$$

Keywords and phrases. Degenerate Cahn–Hilliard equation, single-well potential, continuous Galerkin finite elements, upwind scheme, convergence analysis.

¹ Sorbonne Université, CNRS, Université de Paris, Inria, Laboratoire Jacques-Louis Lions, F-75005 Paris, France.

² MOX, Dipartimento di Matematica, Politecnico di Milano, via Bonardi 9, 20133 Milano, Italy.

[†]Deceased on June 25th 2020.

[‡]Current affiliation: Department of Numerical Analysis and Scientific Computing, Simula Research Laboratory, Oslo, Norway.

*Corresponding author: poulain.alexandre2@gmail.com, poulain@simula.no

and is set in a regular bounded domain $\Omega \subset \mathbb{R}^d$ with $d = 1, 2, 3$. It describes the evolution in time of the (relative) volume fraction $n \equiv n(t, x)$ of one of the two components in a binary mixture. The system is equipped with nonnegative initial data

$$n(0, x) = n^0(x) \in H^1(\Omega), \quad 0 \leq n^0(x) < 1, \quad x \in \Omega,$$

and with zero-flux boundary conditions on the boundary $\partial\Omega$ of Ω

$$\frac{\partial(n - \frac{\sigma}{\gamma}\varphi)}{\partial\nu} = b(n) \frac{\partial(\varphi + \psi'_+(n))}{\partial\nu} = 0, \quad t > 0, \quad x \in \partial\Omega,$$

where ν is the unit normal vector pointing outward $\partial\Omega$.

As pointed at the end of [32], System (1.1) admits a rewriting that can proved to be useful for numerical simulations. Indeed, in the following of this article we use the fact that defining $w = n - \frac{\sigma}{\gamma}\varphi$, System (1.1) can be rewritten as

$$\begin{cases} \frac{\partial n}{\partial t} = \nabla \cdot (b(n) \nabla (\frac{\gamma}{\sigma}n + \psi'_+(n) - \frac{\gamma}{\sigma}w)), \\ -\sigma \Delta w + \frac{\gamma}{\sigma} \psi'_-(w) + w = n, \end{cases} \quad t > 0, \quad x \in \Omega. \quad (1.2)$$

System (1.1) was proposed in [32] as an approximation, in the asymptotic regime whereby the *relaxation parameter* σ vanishes (i.e., $\sigma \rightarrow 0$), of the fourth order Cahn–Hilliard equation [12, 13]. The Cahn–Hilliard (CH) equation describes spinodal decomposition phenomena occurring in binary alloys after quenching: an initially uniform mixed distribution of the alloy undergoes phase separation and a two-phase inhomogeneous structure arises. In its original form, the Cahn–Hilliard equation is written in the form of an evolution equation for n :

$$\partial_t n = \nabla \cdot (b(n) \nabla (\psi'(n) - \gamma \Delta n)), \quad t > 0, \quad x \in \Omega, \quad (1.3)$$

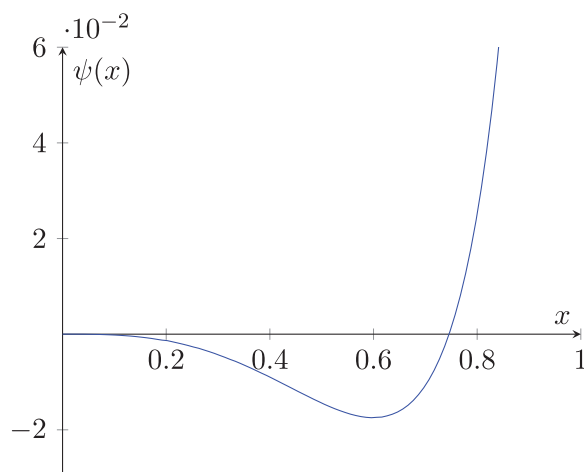
with $n \in [0, 1]$, where the states $n \equiv 0$ and $n \equiv 1$ denote the two pure phases arising after the mixture has undergone the phase separation process. Writing the flux as $\mathbf{J} = -b(n) \nabla \left(\frac{\delta \mathcal{E}[n]}{\delta n} \right)$, equation (1.3) can be interpreted as the conservative gradient flow of the free energy functional

$$\mathcal{E}[n](t) := \int_{\Omega} \left(\frac{\gamma}{2} |\nabla n|^2 + \psi(n) \right) dx.$$

The *homogeneous free energy* ψ describes repulsive and attractive interactions between the two components of the mixture while the regularizing term $\frac{\gamma}{2} |\nabla n|^2$ accounts for partial mixing between the pure phases, leading to a *diffuse interface* separating the states $n \equiv 0$ and $n \equiv 1$, of thickness proportional to $\sqrt{\gamma}$. The parameter $\gamma > 0$ is related to the surface tension at the interface (see, e.g., [30]) and the function b is called *mobility*.

In most of the literature, ψ is a double-well logarithmic potential, often approximated by a smooth polynomial function, with minimums located at the two attraction points that represent pure phases (see, e.g., [18, 22, 23]). The mobility can be either constant [22, 23] or degenerate at the pure phases [6, 21]. We refer to the introductory chapters [20, 31] and to the recent review [30] for an overview of the derivation of the Cahn–Hilliard equation, its analytical properties and its variants.

Recently, the Cahn–Hilliard equation has been considered as a phenomenological model for the description of cancer growth; see, for instance, [3, 17, 37]. In this context, n represents the volume fraction of the tumor in a two-phase mixture containing cancerous cells and a liquid phase, such as water and other nutrients. In biological contexts, a double-well potential appears to be nonphysical. In fact, as suggested by Byrne and Preziosi in [11], a single-well potential of Lennard-Jones type allows for a more suitable description of attractive and repulsive forces acting in the mixture. Following this intuition and building upon previous works [2, 17], in this paper

FIGURE 1. Single-well potential of Lennard-Jones type as in (1.7) with $n^* = 0.6$.

we consider a single-well homogeneous free energy $\psi : [0, 1) \rightarrow \mathbb{R}$, that can be decomposed in a convex and a concave part ψ_{\pm} such that

$$\psi(n) = \psi_+(n) + \psi_-(n), \quad \pm \psi''_{\pm}(n) \geq 0, \quad 0 \leq n < 1. \quad (1.4)$$

Additionally, we consider a singularity at $n = 1$ to represent saturation by one phase (see [11]). Hence, the potential is called *single-well logarithmic* and the singularity is contained in the convex part of the potential. Furthermore, we assume that

$$\psi_+ \in C^2([0, 1)), \quad \psi'_+(1) = \infty, \quad (1.5)$$

and we extend the smooth concave part defined on $[0, 1]$ to \mathbb{R} with

$$\psi_- \in C^3(\mathbb{R}) \quad \psi_-, \psi'_-, \psi''_-, \psi'''_- \quad \text{are bounded and} \quad \frac{\sigma}{\gamma} \|\psi''_-\|_{\infty} < 1. \quad (1.6)$$

The latter assumption is necessary to bound the energy of the system from below. In particular, the example of potential that we have in mind is, for $n \in [0, 1)$,

$$\psi_+(n) = -(1 - n^*) \log(1 - n) - \frac{n^3}{3}, \quad \psi_-(n) = -(1 - n^*) \frac{n^2}{2} - (1 - n^*)n + k, \quad (1.7)$$

where in this case ψ_+ is convex if $n^* \leq 0.7$. In the above form, ψ models cell-cell attraction at small densities ($\psi'(\cdot) < 0$ for $0 < n \leq n^*$ and $\psi'(0) = 0$) and repulsion in overcrowded zones ($\psi'(\cdot) > 0$ for $n \geq n^*$); cf. Figure 1. The quantity $n^* > 0$ represents the value of the cellular density at which repulsive and attractive forces are at equilibrium. With a potential of the form (1.7), the pure phases are represented by the states $n = 0$ and $n = 1$, where $n = 1$ is a singularity for ψ in such a way to avoid overcrowding. In this work, we consider a degenerate mobility $b \in C^1([0, 1]; \mathbb{R}^+)$, such that

$$b(0) = b(1) = 0, \quad b(n) > 0 \quad \text{for} \quad 0 < n < 1. \quad (1.8)$$

Furthermore, the admissible mobility functions that we consider admit the decomposition

$$b(n) = b_1(n)b_2(n),$$

with the extension on \mathbb{R} defined by

$$b_1(n) > 0, \quad \text{for } n > 0, \quad b_1(n) = 0, \quad \text{for } n \leq 0, \quad (1.9)$$

and

$$b_2(n) > 0, \quad \text{for } n < 1, \quad b_2(n) = 0, \quad \text{for } n \geq 1. \quad (1.10)$$

The typical expression in the applications we have in mind is

$$b(n) = n(1 - n)^2. \quad (1.11)$$

Therefore, we can easily see that, considered as transport equations, System (1.1) imposes formally the property $0 \leq n \leq 1$. We also assume an additional property that there is some cancellation at 1 such that

$$b(\cdot)\psi_+''(\cdot) \in C(\mathbb{R}; \mathbb{R}^+). \quad (1.12)$$

For the examples of mobility and potential we described above this assumption is satisfied.

The Cahn–Hilliard equation (1.3) with the logarithmic single-well potential defined in (1.7) and a mobility given by (1.11) has been studied by Agosti *et al.* in [2], where the authors prove well-posedness of the equation for $d \leq 3$. For the relaxed version of the Cahn–Hilliard equation (1.1), well-posedness and convergence to the original Cahn–Hilliard equation as $\sigma \rightarrow 0$ have been proved in [32].

We also mention here some important variants of the Cahn–Hilliard equation that have been used in the context of the modelling of tumors and in which the potential of interaction considered is of double-well form. In particular, Garcke *et al.* [28] proposed a Cahn–Hilliard–Darcy system that models tumor growth and in which the cells can move due to chemotaxis, attractive and repulsive forces. This model also comprises the effect that cells are crawling in a porous medium (*i.e.* the extra-cellular matrix). This effect is represented by a velocity term defined by Darcy’s law. The tumor cells can also proliferate and die depending on the availability of nutrients. Garcke *et al.* [29] extended the model to the multiphase case to study the effect of necrosis. In order to take into account the viscosity of the tissue, and since it is not a valid approximation to consider tissues as a porous medium, Ebenbeck *et al.* [19] proposed a Cahn–Hilliard–Brinkman model in which Brinkman’s law gives the flow velocity this time.

Summary of previous results and specific difficulties. Numerous numerical methods have been developed to solve the Cahn–Hilliard equation (1.3) with smooth and/or logarithmic double-well potential as well as with constant or degenerate mobility. Generally, a numerical scheme for the Cahn–Hilliard equation is evaluated by several aspects: (i) its capacity to keep the energy dissipation (energy stability) and the physical bounds of the solutions; (ii) if it is convergent, and if error bounds can be established; (iii) its efficiency; (iv) its implementation simplicity. To meet the first point concerning the energy stability, several implicit schemes have been proposed. The main drawback of these methods is the necessity to use an iterative method to solve the resulting nonlinear system. To circumvent this issue, unconditionally energy-stable schemes have been proposed based on the splitting of the potential in a convex and a non-convex part. This idea comes from Eyre [26] and leads to unconditionally energy-stable explicit-implicit (*i.e.* semi-implicit) approximations of the model. For references on all the previous numerical methods discussed above, we refer the reader to the review paper [35].

For finite element approximations, most of these results are based on the second-order splitting

$$\begin{cases} \partial_t n = \nabla \cdot (b(n) \nabla \mu), \\ \mu = -\gamma \Delta n + \psi'(n), \end{cases} \quad (1.13)$$

where, μ is called chemical potential; see, *e.g.*, [2, 6, 23].

In [22], Elliot and Songmu propose a finite element Galerkin approximation for the resolution of (1.3) with a smooth double-well potential and constant mobility. The more challenging case of a degenerate mobility

and singular potentials has been considered by Barrett *et al.* in [6], where the authors propose a finite element approximation which employs the second-order splitting (1.13). In particular, the authors provide well-posedness of the finite element approximation as well as a convergence result in the one-dimensional case. Numerical methods to solve the Cahn–Hilliard equation without the splitting technique (1.13) have also been suggested. For instance, in [9] Brenner *et al.* propose a C^0 interior penalty method, a class of discontinuous Galerkin-type approximations.

Even though a single-well potential seems more relevant for biological applications of the Cahn–Hilliard equation, very few works focus on this case. In the already mentioned [2], Agosti and collaborators propose a finite element method to solve equation (1.3) with the homogeneous energy given by (1.7) and a degenerate mobility of the form (1.11). As the authors remark, the main issues arising when considering a single-well logarithmic potential is that the positivity of the solution is not ensured at the discrete level, since the mobility degeneracy set $\{0; 1\}$ does not coincide with the singularity set of the potential, *i.e.*, $n = 1$. Therefore, the absence of cells represents an unstable equilibrium of the potential. In [2], the authors design a finite element scheme which preserves positivity by the means of a discrete variational inequality, as also suggested in [6]. More recently, in [1], Agosti has presented a discontinuous Galerkin finite element discretization of the equation where, again, the positivity of the discrete solution is ensured thanks to a discrete variational inequality.

In this paper, we take advantage of the equivalent rewriting (1.2) of the RDCH system and use previous results obtained on the analysis of finite element schemes derived for degenerate parabolic equations. Namely, we use results obtained by Cancès and Guichard [14] for a Control-Volume-Finite-Element (CVFE) used to simulate the anisotropic porous medium equation. In this work, the nonnegativity of the discrete solution is achieved by a suitable definition of the convective mobility that is discretized on control volumes. The compactness of time and space translates are achieved using discrete energy estimates, and the convergence analysis is obtained using the Frechet–Kolmogorov theorem. This work is the basis of a subsequent paper by Cancès *et al.* [15] in which a CVFE scheme for an anisotropic Keller–Segel system is analyzed. Therefore, using the fact that the RDCH system is close to the Keller–Segel model (*i.e.* they are both composed of one degenerate parabolic equation and one elliptic equation), we combine the results of the previously presented works on CVFE schemes to perform our analysis.

Contents of the paper. The aim of this paper is to present and analyze two finite element approximations of the relaxed Cahn–Hilliard equation with single-well potential (1.7) and degenerate mobility (1.11) in dimensions $d = 1, 2, 3$. More in detail, we propose two different time discretizations leading to one nonlinear scheme on which we can prove analytically some important properties and one efficient linear scheme that retrieve these important properties during numerical simulations (but we can not prove them analytically). However, for the nonlinear scheme, we prove analytically: (i) well-posedness of the numerical approximation; (ii) nonnegativity of discrete solutions ensured by adapting the upwind technique to the finite element approximation method; (iii) dissipation of a discrete energy; (iv) convergence of discrete solutions.

The main novelty of our work is to propose an alternative to the second-order splitting (1.13) by replacing the chemical potential μ by its relaxed approximation φ , solution to a second order elliptic equation with diffusivity $0 < \sigma \ll \gamma$. The relaxed system is based on the analysis performed in [32], where the authors prove well-posedness of the system as well as the convergence, as $\sigma \rightarrow 0$, of weak solutions of (1.1) to the ones of the original Cahn–Hilliard equation (1.3). For the analysis that follows, it is worth noticing that System (1.1) admits the energy functional

$$\mathcal{E}_\sigma[n](t) := \int_{\Omega} \left\{ \frac{\gamma}{2} \left| \nabla \left(n - \frac{\sigma}{\gamma} \varphi \right) \right|^2 + \frac{\sigma}{2\gamma} |\varphi|^2 + \psi_+(n) + \psi_- \left(n - \frac{\sigma}{\gamma} \varphi \right) \right\} dx, \quad (1.14)$$

that, as proved in [32], is decreasing in time, *i.e.*,

$$\frac{d\mathcal{E}_\sigma[n]}{dt} = - \int_{\Omega} b(n) |\nabla (\varphi + \psi'_+(n))|^2 dx \leq 0, \quad t > 0.$$

We also notice that the convex/concave splitting of ψ is different from the one employed, *e.g.*, in [2] and is motivated by the need to retrieve energy dissipation as well as by the fact that we can take advantage of the linearity of ψ'_- to achieve regularity results on w . Furthermore, we observe that the relaxed Cahn–Hilliard system bears some similarities with the Keller–Segel model with additional cross diffusion, proposed and analyzed in [7, 16].

In this work, we use another definition of the continuous solutions for the equivalent system (1.2). To do so, we define the non-decreasing continuous function $\eta : \mathbb{R} \rightarrow \mathbb{R}$ by

$$\eta(r) = \int_0^r \sqrt{b(s)} \, ds, \quad \forall r \in \mathbb{R}. \quad (1.15)$$

Similarly, we also define the non-decreasing continuous function $\zeta : \mathbb{R} \rightarrow \mathbb{R}$, using the properties of $b(s)\psi''_+(s)$ stated in (1.12), such that

$$\zeta(r) = \int_0^r \sqrt{b(s)\psi''_+(s)} \, ds, \quad \forall r \in \mathbb{R}. \quad (1.16)$$

With the definition of these new functions, we define the solutions of the RDCH system (1.2).

Definition 1.1 (Weak solutions of RDCH). The functions (n, w) such that

$$\begin{cases} 0 \leq n \leq 1, \text{ a.e. in } \Omega_T, \\ \eta(n) \in L^2(0, T; H^1(\Omega)), \quad \zeta(n) \in L^2(0, T; H^1(\Omega)), \\ w \in L^\infty(\Omega_T) \cap L^2(0, T; H^1(\Omega)), \end{cases} \quad (1.17)$$

are the weak solutions of the relaxed-degenerate Cahn–Hilliard model (1.2) in the following sense: For all test functions $\chi_1, \chi_2 \in C_c^\infty(\Omega_T, \mathbb{R}^+)$ with $\chi_1(T, \cdot) = \chi_2(T, \cdot) = 0$, we have

$$-\int_\Omega n^0 \chi_1 - \int_{\Omega_T} n \frac{\partial \chi_1}{\partial t} + \frac{\gamma}{\sigma} \int_{\Omega_T} \sqrt{b(n)} \nabla \eta(n) \nabla \chi_1 + \int_{\Omega_T} \sqrt{b(n)\psi''_+(n)} \nabla \zeta(n) \nabla \chi_1 = \frac{\gamma}{\sigma} \int_{\Omega_T} b(n) \nabla w \nabla \chi_1, \quad (1.18)$$

$$\sigma \int_{\Omega_T} \nabla w \nabla \chi_2 + \int_{\Omega_T} \left(\frac{\gamma}{\sigma} \psi'_-(w) + w \right) \chi_2 = \int_{\Omega_T} n \chi_2. \quad (1.19)$$

This paper is organized as follows. Section 2 presents the details of the Control-Volume-Finite-Element framework and the assumptions on the mesh. In Section 3, we introduce a nonlinear semi-implicit finite element approximation of the relaxed Cahn–Hilliard equation (1.1). The definition of the upwind mobility coefficient is given using a constant approximation of the mobility term on specific control volumes. Sections 3.2 and 3.3 are dedicated to the presentation of the properties of this nonlinear scheme, such as the nonnegativity of the discrete solutions, mass conservation, and *a priori* estimates. The existence of discrete solutions to this problem is given in Section 3.4. In Sections 3.5 and 3.6, we derive compactness estimates and prove the convergence of the discrete solutions to the weak solutions of the continuous relaxed Cahn–Hilliard model defined in Definition 1.1. Then, Section 4 is devoted to the description of an efficient linear semi-implicit scheme. The existence and the nonnegativity of a unique solution are given for a suitable upwind approximation of the mobility coefficient. Finally, in Section 5, we present numerical simulations using our linear semi-implicit scheme in one and two dimensions. These numerical simulations are in good agreement with previous numerical results obtained for the degenerate Cahn–Hilliard equation with single-well logarithmic potential.

2. NOTATIONS AND ASSUMPTIONS

We first set up the notations we will use in the numerical discretization and recall some well-known properties we employ in the analysis.

Geometric and functional setting. Let $\Omega \subset \mathbb{R}^d$ with $d = 1, 2, 3$ be a polyhedral domain. We indicate the usual Lebesgue and Sobolev spaces by respectively $L^p(\Omega)$, $W^{m,p}(\Omega)$ with $H^m(\Omega) := W^{m,2}(\Omega)$, where $1 \leq p \leq +\infty$ and $m \in \mathbb{N}$. We denote the corresponding norms by $\|\cdot\|_{m,p,\Omega}$, $\|\cdot\|_{m,\Omega}$ and semi-norms by $|\cdot|_{m,p,\Omega}$, $|\cdot|_{m,\Omega}$. The standard L^2 inner product will be denoted by $(\cdot, \cdot)_\Omega$ and the duality pairing between $(H^1(\Omega))'$ and $H^1(\Omega)$ by $\langle \cdot, \cdot \rangle_\Omega$.

Let \mathcal{T}^h , $h > 0$ be a conformal mesh on the domain Ω which is defined by N_{el} disjoint piecewise linear mesh elements, denoted by $K \in \mathcal{T}^h$, such that $\bar{\Omega} = \bigcup_{K \in \mathcal{T}^h} \bar{K}$. The elements are triangles for $d = 2$ and tetrahedra for $d = 3$. We let $h := \max_K h_K$ refers to the level of refinement of the mesh, where $h_K := \text{diam}(K)$ for $K \in \mathcal{T}^h$. We define by κ_K the minimal perpendicular length of K and $\kappa_h = \min_{K \in \mathcal{T}^h} \kappa_K$. We assume that the mesh is quasi-uniform, *i.e.*, it is shape-regular and there exists a constant $C > 0$ such that

$$h_K \geq Ch, \quad \forall K \in \mathcal{T}^h.$$

Moreover, we assume that the mesh is acute, *i.e.*, for $d = 2$ the angles of the triangles do not exceed $\frac{\pi}{2}$ and for $d = 3$ the angle between two faces of the same tetrahedron do not exceed $\frac{\pi}{2}$.

We define the set of nodal points $J_h = \{x_j\}_{j=1,\dots,N_h}$, where the number of nodes is $N_h := \text{card}(J_h)$, and we assume that each x_j is a vertex of a simplex $K \in \mathcal{T}^h$. We denote by Λ_i the set of nodes connected to the node x_i by an edge and G_h the maximal number of connected nodes, *i.e.* $G_h = \max_{x_i \in J_h} \Lambda_i$. For two nodes x_i, x_j , if they are connected by an edge, we denote this latter e_{ij} .

We also define the barycentric dual mesh associated to \mathcal{T}^h . On each element $K \in \mathcal{T}^h$, the barycentric coordinates of an arbitrary point $x \in K$ are defined by the real numbers λ_i with $i = 1, \dots, n_K$ such that

$$\sum_{i=1}^{n_K} \lambda_i = 1, \quad \text{and} \quad x = \sum_{i=1}^{n_K} \lambda_i P_i,$$

where n_K is the number of nodes of the element K . We define the barycentric subdomains associated to the vertex $P_i \in K_k$ (where K_k refers to the k -th element of \mathcal{T}^h and $k = 1, \dots, N_{\text{el}}$), by

$$D_i^k := \bigcap_{\substack{j=1 \\ j \neq i}}^{n_K} \{x; x \in K_k \text{ and } \lambda_j(x) \leq \lambda_i(x)\}.$$

Therefore, for each node of the mesh \mathcal{T}^h , we define the barycentric cell of the dual mesh

$$D_i := \bigcup_k \{D_i^k; K_k \in \mathcal{T}^h \text{ such that } x_i \in K_k\}.$$

We also add another subdivision of the mesh that relies on the definition of the barycentric cells and that will be useful to define the upwind approximation of the mobility. Using the barycentric coordinates, we subdivide each element $K \in \mathcal{T}^h$ in $d + 1$ subdomains, such that, for $i = 1, 2, 3$, $j = 2, 3$, and $i \neq j$, we have the subdomain

$$\tilde{D}_{ij}^K = \{x \in K | \lambda_i, \lambda_j \geq \lambda_k, \quad k \neq i, j\}.$$

In the following of the article, we use the terminology *Diamond cell* for the two \tilde{D}_{ij}^K that share the edge e_{ij} .

To illustrate what are these subdomains, we represent graphically what is \tilde{D}_{ij}^T on Figure 2 as well as what is a diamond cell for $d = 2$.

We introduce the set of piecewise linear functions $\chi_j \in C(\bar{\Omega})$ associated with the nodal point $x_j \in J_h$, that satisfies $\chi_j(x_i) = \delta_{ij}$, where δ_{ij} is the Kronecker's delta. We introduce the \mathbb{P}^1 conformal finite element space V^h

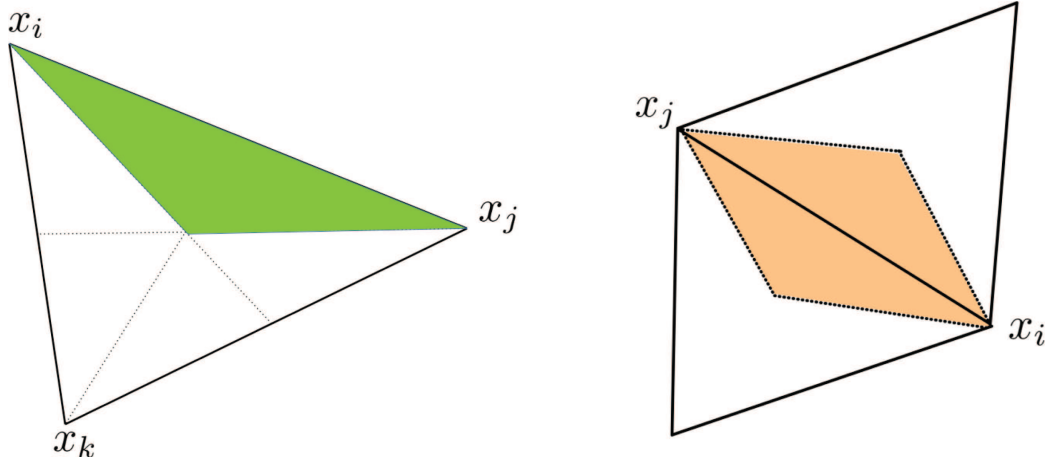


FIGURE 2. Illustration of the subdomain \tilde{D}_{ij}^K (left, green) and of a diamond cell (right, orange) for $d = 2$.

associated with \mathcal{T}^h , where $\mathbb{P}^1(K)$ denotes the space of polynomials of order 1 on K :

$$V^h := \{\chi \in C(\overline{\Omega}) : \chi|_K \in \mathbb{P}^1(K), \quad \forall K \in \mathcal{T}^h\} \subset H^1(\Omega).$$

Furthermore, we let K^h be the subset containing the nonnegative elements of V^h , namely

$$K^h := \{\chi \in V^h : \chi \geq 0 \quad \text{in } \Omega\}.$$

We denote by $\pi^h : C(\overline{\Omega}) \rightarrow V^h$ the Lagrange interpolation operator corresponding to the discretized domain \mathcal{T}^h , defined as

$$\pi^h f(x) = \sum_{j=1}^{N_h} f(x_j) \chi_j(x), \quad f \in C(\overline{\Omega}).$$

We also use the lumped space \hat{V}_h defined by

$$\hat{V}^h := \{\hat{\chi} : \text{piecewise constant over barycentric domains, i.e. } \hat{\chi}(x) = \hat{\chi}(x_i), \forall x \in D_i\}.$$

Defining $\hat{\chi}_i \in L^\infty(\Omega)$ the characteristic function of the barycentric domain D_i associated with each node x_i (for $i = 1, \dots, N_h$) of the mesh, we easily see that $\{\hat{\chi}_j\}_{j=1, \dots, N_h}$ forms a basis of \hat{V}^h . Adding the property $\hat{\chi}_i(x_j) = \delta_{ij}$, we see that the two basis $\{\hat{\chi}_j\}_{j=1, \dots, N_h}$ and $\{\chi_j\}_{j=1, \dots, N_h}$ are associative, i.e. $\chi(x_i) = \hat{\chi}(x_i)$ for all $x_i \in J_h$. Therefore, we also define the lumped operator $\hat{\pi}^h : C(\overline{\Omega}) \rightarrow \hat{V}^h$ by

$$\hat{\pi}^h f(x) = \sum_{j=1}^{N_h} f(x_j) \hat{\chi}_j(x), \quad f \in C(\overline{\Omega}).$$

On $C(\overline{\Omega})$ we define the approximate scalar product as

$$(f, g)^h := \int_{\Omega} \pi^h(f(x)g(x)) \, dx, \quad f, g \in C(\overline{\Omega}).$$

Furthermore, since $\forall f, g \in C(\overline{\Omega})$, we have

$$\begin{aligned}
\int_{\Omega} \pi^h(f(x)g(x)) \, dx &= \sum_{K \in \mathcal{T}^h} \int_K \pi^h(f(x)g(x)) \, dx, \\
&= \frac{1}{d+1} \sum_{K \in \mathcal{T}^h} |K| \sum_{x_i \in K} f(x_i)g(x_i), \\
&= \int_{\Omega} \hat{\pi}^h(f(x)g(x)) \, dx,
\end{aligned}$$

where $x_i \in K$ denotes the vertices of the element K . We denote the corresponding discrete semi-norm as $|\cdot|_h = \left[(\cdot, \cdot)^h \right]^{\frac{1}{2}}$.

We denote by $P_h^0 : L^2(\Omega) \rightarrow V^h$ the L^2 projection operator and by $\hat{P}_h^0 : L^2(\Omega) \rightarrow V^h$ its lumped version, defined by

$$\begin{aligned}
(P_h^0 v, \chi) &= (v, \chi) \quad \forall v \in L^2(\Omega) \quad \text{and} \quad \forall \chi \in V^h, \\
(\hat{P}_h^0 v, \chi)^h &= (v, \chi) \quad \forall v \in L^2(\Omega) \quad \text{and} \quad \forall \chi \in V^h.
\end{aligned}$$

Inequalities. We summarize important inequalities that will be used later on for the analysis of the numerical schemes. We will use the following inequalities (see, *e.g.*, [33]):

$$|\chi|_{m,p_2} \leq Ch^{-d(\frac{1}{p_1} - \frac{1}{p_2})} |\chi|_{m,p_1} \quad \forall \chi \in V^h, \quad 1 \leq p_1 \leq p_2 \leq +\infty, \quad m = 0, 1; \quad (2.1)$$

$$\|\chi\|_0^2 \leq (\chi, \chi)^h \leq (d+2) \|\chi\|_0^2 \quad \forall \chi \in V^h. \quad (2.2)$$

Concerning the interpolation operator, the following result holds [8]:

$$\lim_{h \rightarrow 0} \|v - \pi^h(v)\|_{0,\infty} = 0 \quad \forall v \in C(\overline{\Omega}), \quad (2.3)$$

and we have [34],

$$|(v, \eta)^h - (v, \eta)| \leq Ch^2 \|\nabla v\|_0 \|\nabla \eta\|_0, \quad v, \eta \in V^h. \quad (2.4)$$

For the L^2 projection operator, the following inequality holds [8]

$$\|v - P_h^0 v\|_0 + h |v - P_h^0 v|_1 \leq Ch^m |v|_m \quad v \in H^m(\Omega), \quad m = 1, 2. \quad (2.5)$$

Finite element matrices. We define M and Q respectively the mass and stiffness matrix. M_l is the lumped mass matrix, that is the diagonal matrix where each coefficient is the sum of the associated row of M

$$\begin{aligned}
M_{ij} &= \int_{\Omega} \chi_i \chi_j \, dx, & \text{for } i, j = 1, \dots, N_h, \\
Q_{ij} &= \int_{\Omega} \nabla \chi_i \cdot \nabla \chi_j \, dx, & \text{for } i, j = 1, \dots, N_h, \\
M_{l,ii} &:= \sum_{j=1}^{N_h} M_{ij}, & \text{for } i, j = 1, \dots, N_h.
\end{aligned}$$

We recall some important properties of the stiffness matrix. We start by

$$\sum_{j \neq i} Q_{ij} = Q_{ii}. \quad (2.6)$$

Furthermore, from the fact that the triangulation is of acute type, we know from [27] that

$$Q_{ij} \leq 0 \quad \text{for } i \neq j, \quad \text{and} \quad 0 < \frac{|Q_{ij}|}{M_{l,ii}} \leq \frac{Q_{ii}}{M_{l,ii}} \leq \frac{(d+1)}{\kappa_h^2}. \quad (2.7)$$

3. NONLINEAR CVFE SCHEME

3.1. Description of the scheme

Given $N_T \in \mathbb{N}$, let $\Delta t := T/N_T$ be the constant time step size and $t^k := k\Delta t$, for $k = 0, \dots, N_T - 1$. We consider a partitioning of the time interval $[0, T] = \bigcup_{k=0}^{N_T-1} [t^k, t^{k+1}]$ with $0 < T < +\infty$. We approximate the continuous time derivative by $\frac{\partial n_h}{\partial t} \approx \frac{n_h^{k+1} - n_h^k}{\Delta t}$. Using the finite element approximations of n and w ,

$$n_h^k(x) := \sum_{j=1}^{N_h} n_j^k \chi_j(x), \quad \text{and} \quad w_h^k(x) := \sum_{j=1}^{N_h} w_j^k \chi_j(x),$$

where $\{n_j^k\}_{j=1, \dots, N_h}$ and $\{w_j^k\}_{j=1, \dots, N_h}$ are the unknown degrees of freedom, we introduce the following finite element approximation of System (1.1).

Problem P. For each $k = 0, \dots, N_T - 1$, find $\{n_h^{k+1}, w_h^{k+1}\}$ in $K^h \times V^h$ such that $\forall \chi_1, \chi_2 \in V^h$

$$\left(\frac{n_h^{k+1} - n_h^k}{\Delta t}, \chi_1 \right)^h + \left(\left(\frac{\gamma}{\sigma} \hat{B}(n_h^{k+1}) + \widehat{(B\psi_+'')}(n_h^{k+1}) \right) \nabla n_h^{k+1}, \nabla \chi_1 \right) = \frac{\gamma}{\sigma} \left(\tilde{B}(n_h^{k+1}) \nabla w_h^{k+1}, \nabla \chi_1 \right), \quad (3.1)$$

$$\sigma \left(\nabla w_h^{k+1}, \nabla \chi_2 \right) + \left(w_h^{k+1} + \frac{\gamma}{\sigma} \psi'_-(w_h^k), \chi_2 \right)^h = (n_h^k, \chi_2)^h, \quad (3.2)$$

where $\tilde{B}(\cdot)$ is the discrete upwind approximation of the continuous mobility $b(\cdot)$ for the convective term while $\hat{B}(\cdot)$ and $\widehat{(B\psi_+'')}(\cdot)$ are the constant approximations of the mobility and second derivative of the convex part of the potential multiplied by the mobility. These latter quantities are defined in equations (3.11)–(3.13) respectively. The initial condition for the discrete problem is chosen such that for $n^0(x) \in H^1(\Omega)$, we have

$$\begin{cases} n_h^0 = \pi^h(n^0(x)), & \text{for } d = 1, \\ n_h^0 = \hat{P}_h^0(n^0(x)), & \text{for } d = 2, 3, \end{cases} \quad (3.3)$$

and w_h^0 is the unique solution of

$$\sigma \left(\nabla w_h^0, \nabla \chi \right) + \left(w_h^0 + \frac{\gamma}{\sigma} \psi'_-(w_h^0), \chi \right)^h = (n_h^0, \chi)^h, \quad \forall \chi \in V^h. \quad (3.4)$$

To preserve the nonnegativity of the discrete solutions, we compute the mobility coefficient employing an implicit upwind method adapted to the finite element setting. The explanation on how to adapt the upwind method requires us to define the matrix formulation of the problem (3.1) and (3.2).

Matrix form. For $k = 0, \dots, N_T - 1$, let \underline{n}^k and \underline{w}^k be the vectors

$$\underline{n}^k := [n_1^k, \dots, n_{N_h}^k]^T, \quad \underline{w}^k := [w_1^k, \dots, w_{N_h}^k]^T.$$

We can then rewrite System (3.1) and (3.2) in its matrix form

$$M_l \underline{n}^{k+1} + \Delta t R^{k+1} \underline{n}^{k+1} = M_l \underline{n}^k + \Delta t U^{k+1} \underline{w}^{k+1}, \quad (3.5)$$

$$(\sigma Q + M_l) \underline{w}^{k+1} + \frac{\gamma}{\sigma} M_l \underline{\psi}'_- = M_l \underline{n}^k, \quad (3.6)$$

where $\underline{\psi}'_-$ is the vector containing the nodal values

$$\left(\underline{\psi}'_- \right)_i = \psi'_-(w_i^k) \quad i = 1, \dots, N_h.$$

In the above matrix form, we have used the definitions of the finite element matrices associated with $U^{k+1} = U(n_h^{k+1}) = \left(\tilde{B}(n_h^{k+1}) \nabla \cdot, \nabla \cdot \right)$, such that

$$U_{ij}^{k+1} = \frac{\gamma}{\sigma} \int_{\Omega} \tilde{B}(n_h^{k+1}) \nabla \chi_i \nabla \chi_j \, dx = \frac{\gamma}{\sigma} \int_{\Omega} \tilde{B}_{ij}^{k+1} \nabla \chi_i \cdot \nabla \chi_j \, dx, \quad \text{for } i, j = 1, \dots, N_h, \quad (3.7)$$

and for $R^{k+1} = R(n_h^{k+1}) = \left(\frac{\gamma}{\sigma} \hat{B}(n_h^{k+1}) + \widehat{(B\psi_+'')}(n_h^{k+1}) \nabla \cdot, \nabla \cdot \right)$, we have

$$\begin{aligned} R_{ij}^{k+1} &= \int_{\Omega} \frac{\gamma}{\sigma} \hat{B}(n_h^{k+1}) + \widehat{(B\psi_+'')}(n_h^{k+1}) \nabla \chi_i \cdot \nabla \chi_j \, dx \\ &= \int_{\Omega} \frac{\gamma}{\sigma} \hat{B}_{ij}^{k+1} + \widehat{(B\psi_+'')}_{ij}^{k+1} \nabla \chi_i \cdot \nabla \chi_j \, dx, \quad \text{for } i, j = 1, \dots, N_h. \end{aligned} \quad (3.8)$$

From (2.6) and (2.7), we can interpret the first equation of System (3.1) and (3.2) as using the finite volume method *i.e.*, for every $x_i \in J_h$, we have

$$\begin{aligned} M_{l,ii} n_i^{k+1} + \Delta t \sum_{x_j \in \Lambda_i} \left(\frac{\gamma}{\sigma} \hat{B}_{ij}^{k+1} + \widehat{(B\psi_+'')}_{ij}^{k+1} \right) |Q_{ij}| (n_i^{k+1} - n_j^{k+1}) \\ = M_{l,ii} n_i^k + \frac{\gamma \Delta t}{\sigma} \sum_{x_j \in \Lambda_i} \tilde{B}_{ij}^{k+1} |Q_{ij}| (w_i^{k+1} - w_j^{k+1}). \end{aligned} \quad (3.9)$$

Therefore, defining $F_{ij}^{k+1} = H_{ij}^{k+1} - G_{ij}^{k+1}$ with

$$\begin{aligned} H_{ij}^{k+1} &= H(n_i^{k+1}, n_j^{k+1}, (\delta n^{k+1})_{ij}) = \left(\frac{\gamma}{\sigma} \hat{B}_{ij}^{k+1} + \widehat{(B\psi_+'')}_{ij}^{k+1} \right) Q_{ij} (-(\delta n^{k+1})_{ij}), \\ G_{ij}^{k+1} &= G(n_i^{k+1}, n_j^{k+1}, (\delta w^{k+1})_{ij}) = \frac{\gamma}{\sigma} \tilde{B}_{ij}^{k+1} Q_{ij} (-(\delta w^{k+1})_{ij}), \end{aligned}$$

where $(\delta w)_{ij}$ denotes the difference $(w_i^{k+1} - w_j^{k+1})$. Thus, the scheme reads

$$\begin{cases} F_{ij}^{k+1} + F_{ji}^{k+1} = 0, \end{cases} \quad (3.10a)$$

$$\begin{cases} M_{l,ii} n_i^{k+1} = M_{l,ii} n_i^k - \Delta t \sum_{j \neq i} F_{ij}^{k+1}, \quad \forall x_i \in J_h, \end{cases} \quad (3.10b)$$

$$\begin{cases} \sigma (\nabla w_h^{k+1}, \nabla \chi) + \left(w_h^{k+1} + \frac{\gamma}{\sigma} \psi'_-(w_h^k), \chi \right)^h = (n_h^k, \chi)^h. \end{cases} \quad (3.10c)$$

Let us now describe in detail how the coefficients \tilde{B}_{ij}^{k+1} , \hat{B}_{ij}^{k+1} and $\widehat{(B\psi_+'')}_{ij}^{k+1}$ are computed.

Constant approximation of mobility and second derivative of potential. We define the upwind mobility coefficient associated with the convective term by

$$\tilde{B}_{ij}^{k+1} := \begin{cases} b_1(n_i^{k+1}) b_2(n_j^{k+1}), & \text{if } w_j^{k+1} - w_i^{k+1} > 0, \\ b_1(n_j^{k+1}) b_2(n_i^{k+1}), & \text{otherwise,} \end{cases} \quad i, j = 1, \dots, N_h. \quad (3.11)$$

The mobility function and second derivative of the convex part of the potential associated to the diffusion term are defined by

$$\hat{B}_{ij}^{k+1} := \max_{[n_i^{k+1}, n_j^{k+1}]} b(s) \quad i, j = 1, \dots, N_h, \quad (3.12)$$

and

$$\widehat{(B\psi''_+)}_{ij}^{k+1} := \max_{[n_i^{k+1}, n_j^{k+1}]} b(s)\psi''_+(s) \quad i, j = 1, \dots, N_h. \quad (3.13)$$

The approximation (3.11) is similar to the one used in [4] for the one-dimensional finite volume discretization of the Keller–Segel system. Furthermore, our adaptation of the upwind method in the finite element context is also close in spirit to the one proposed by Baba and Tabata in [5], where the authors also used barycentric coordinates to define the basis functions.

We remark that the coefficients \tilde{B}_{ij}^{k+1} , \hat{B}_{ij}^{k+1} and $\widehat{(B\psi''_+)}_{ij}^{k+1}$ are constant and uniquely defined along each edge of the mesh (*i.e.* in the diamond cells). Yet, this method is well suited for a standard assembling procedure and, as a result, is simpler to implement in already existing finite element software since it requires only the adaptation of the calculation of a non-constant matrix. This method can also be adapted for the simulation of other advection-diffusion equations to preserve the nonnegativity of solutions.

From this approximation of the mobility we have the following properties concerning the convective flux.

Proposition 3.1 (Properties of the convective flux). *For any $(a, b, c) \in \mathbb{R}^3$, we have:*

- (1) $G(\cdot, b, c)$ is non-decreasing, and $G(a, \cdot, c)$ is non-increasing;
- (2) $G(a, b, c) = -G(b, a, -c)$;
- (3) $G(a, a, c) = b(a)(-c)$;
- (4) there is a constant $C > 0$ such that $|G(a, b, c)| \leq C(|a| + |b|)|c|$;
- (5) there is a modulus of continuity $\omega : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ such that for all $(a', b', c) \in \mathbb{R}^3$

$$|G(a, b, c) - G(a', b', c')| \leq |c| \omega(|a - a'| + |b - b'|).$$

3.2. Discrete maximum principle, non-negativity, and mass conservation

Proposition 3.2 (Non-negativity of n_h^{k+1} and upper bound). *The numerical scheme (3.1) and (3.2) preserves the non-negativity of the solution, i.e. for all $k = 0, \dots, N_T - 1$, we have $0 \leq n_h^{k+1} \leq 1$.*

Proof. Step 1. Non-negativity. We prove the non-negativity by induction. We assume that $\forall x_i \in J_h, 0 \leq n_i^k \leq 1$, and $n_i^{k+1} = \min_{x_j \in J_h} (n_j^{k+1})$. Then, we multiply equation (3.10b) by $-(n_i^{k+1})^- = -\max(0, -n_i^{k+1})$ to obtain

$$\begin{aligned} -\sum_{j=1}^{N_h} (\hat{\chi}_j, \hat{\chi}_i) \left(\frac{n_i^{k+1} - n_i^k}{\Delta t} \right) (n_i^{k+1})^- &= \sum_{j \neq i} \left(\frac{\gamma}{\sigma} \hat{B}_{ij}^{k+1} + \widehat{(B\psi''_+)}_{ij}^{k+1} \right) Q_{ij} (n_j^{k+1} - n_i^{k+1}) (n_i^{k+1})^- \\ &\quad - \frac{\gamma}{\sigma} \sum_{j \neq i} \tilde{B}_{ij}^{k+1} Q_{ij} (w_j^{k+1} - w_i^{k+1}) (n_i^{k+1})^-. \end{aligned}$$

From the non-negativity of \hat{B}_{ij}^{k+1} , and $\widehat{(B\psi''_+)}_{ij}^{k+1}$ as well as (2.7), we know that

$$\sum_{j \neq i} \left(\frac{\gamma}{\sigma} \hat{B}_{ij}^{k+1} + \widehat{(B\psi''_+)}_{ij}^{k+1} \right) Q_{ij} (n_j^{k+1} - n_i^{k+1}) (n_i^{k+1})^- \leq 0.$$

Then, if $w_j^{k+1} - w_i^{k+1} > 0$, $\tilde{B}_{ij}^{k+1} = b_1(n_i^{k+1})b_2(n_j^{k+1})$, and from the extension of $b_1(s)$ by zero for $s < 0$, we have

$$-\frac{\gamma}{\sigma} \sum_{j \neq i} \tilde{B}_{ij}^{k+1} Q_{ij} (w_j^{k+1} - w_i^{k+1}) (n_i^{k+1})^- \leq 0.$$

Altogether, we obtain

$$-\sum_{j=1}^{N_h} (\hat{\chi}_j, \hat{\chi}_i) \left(\frac{n_i^{k+1} - n_i^k}{\Delta t} \right) (n_i^{k+1})^- \leq 0,$$

and, hence, $(n_i^{k+1})^- = 0$ which implies $n_i^{k+1} \geq 0$ and $n_h^{k+1} \geq 0$.

Step 2. Upper bound. To prove the upper bound we repeat the previous argument but this time we assume $\forall x_i \in J_h$, $0 \leq n_i^k \leq 1$, and $n_i^{k+1} = \max_{x_j \in J_h} (n_j^{k+1})$. We multiply equation (3.10b) by $(n_i^{k+1} - 1)^+ = \max(0, n_i^{k+1} - 1)$ to obtain

$$\begin{aligned} \sum_{j=1}^{N_h} (\hat{\chi}_j, \hat{\chi}_i) \left(\frac{n_i^{k+1} - n_i^k}{\Delta t} \right) (n_i^{k+1} - 1)^+ &= - \sum_{j \neq i} \left(\frac{\gamma}{\sigma} \hat{B}_{ij}^{k+1} + \widehat{(B\psi_+'')_{ij}}^{k+1} \right) Q_{ij} (n_j^{k+1} - n_i^{k+1}) (n_i^{k+1} - 1)^+ \\ &\quad + \frac{\gamma}{\sigma} \sum_{j \neq i} \tilde{B}_{ij}^{k+1} Q_{ij} (w_j^{k+1} - w_i^{k+1}) (n_i^{k+1} - 1)^+. \end{aligned}$$

Since $n_i^{k+1} = \max_{x_j \in J_h} (n_j^{k+1})$ and from the non-negativity of \hat{B}_{ij}^{k+1} and $\widehat{(B\psi_+'')_{ij}}^{k+1}$ as well as (2.7), we have

$$- \sum_{j \neq i} \left(\frac{\gamma}{\sigma} \hat{B}_{ij}^{k+1} + \widehat{(B\psi_+'')_{ij}}^{k+1} \right) Q_{ij} (n_j^{k+1} - n_i^{k+1}) (n_i^{k+1} - 1)^+ \leq 0.$$

Then, if $w_j^{k+1} - w_i^{k+1} \leq 0$, we have $\tilde{B}_{ij}^{k+1} = b_1(n_j^{k+1})b_2(n_i^{k+1})$, and from the extension of $b_2(s)$ by zero for $s > 1$, we obtain

$$\frac{\gamma}{\sigma} \sum_{j \neq i} \tilde{B}_{ij}^{k+1} Q_{ij} (w_j^{k+1} - w_i^{k+1}) (n_i^{k+1} - 1)^+ \leq 0.$$

Altogether, we have

$$\sum_{j=1}^{N_h} (\hat{\chi}_j, \hat{\chi}_i) \left(\frac{n_i^{k+1} - n_i^k}{\Delta t} \right) (n_i^{k+1} - 1)^+ \leq 0,$$

which implies $(n_i^{k+1} - 1)^+ = 0$, and, hence, $n_i^{k+1} \leq 1$. Therefore, we obtain the result. \square

The previous result allows to show that w_h^{k+1} is confined in a threshold $[-C, C]$ with $C > 0$ finite.

Proposition 3.3. *The numerical scheme (3.1) and (3.2) preserves a lower and upper bound for w_h^{k+1} , i.e. for all $k = 0, \dots, N_T - 1$, we have $-C \leq w_h^{k+1} \leq C$, for $C > 0$ finite.*

Proof. The result is found using equation (3.2), Proposition 3.2, and the boundedness of $\psi'_-(s)$ for all $s \in \mathbb{R}$ (given by (1.6)). \square

Proposition 3.4 (Conservation of mass). *The finite element numerical scheme (3.1) and (3.2) preserves the initial mass, i.e. for all $k = 0, \dots, N_T - 1$, we have*

$$\int_{\Omega} n_h^0 \, dx = \int_{\Omega} n_h^{k+1} \, dx.$$

Proof. To prove mass conservation, we use the identity, for each $x_i \in J_h$

$$\sum_{x_j \in \Lambda_i} |U_{ij}^{k+1}| = U_{ii}^{k+1}, \quad \text{and} \quad \sum_{x_j \in \Lambda_i} |R_{ij}^{k+1}| = R_{ii}^{k+1}. \quad (3.14)$$

Summing over the nodes in equation (3.1), we obtain

$$\begin{aligned} \sum_{i=1}^{N_h} \sum_{j=1}^{N_h} (\hat{\chi}_j, \hat{\chi}_i) (n_h^{k+1} - n_h^k)(x_i) &= - \Delta t \sum_{i=1}^{N_h} \sum_{j=1}^{N_h} U_{ij}^{k+1} (-w_h^{k+1})(x_j) \\ &\quad - \Delta t \sum_{i=1}^{N_h} \sum_{j=1}^{N_h} R_{ij}^{k+1} n_h^{k+1}(x_j). \end{aligned}$$

Using the symmetry of the matrices U^{k+1} and R^{k+1} , the property (3.14), we obtain

$$\sum_{i=1}^{N_h} \sum_{j=1}^{N_h} (\hat{\chi}_j, \hat{\chi}_i) (n_h^{k+1} - n_h^k)(x_i) = 0,$$

which implies mass conservation. \square

3.3. Energy stability and *a priori* estimates

The finite element numerical scheme (3.1) and (3.2) preserves the dissipation of the energy at the discrete level.

Proposition 3.5 (Energy estimate for w_h^{k+1}). *System (3.1) and (3.2) admits the following a priori estimate*

$$\sum_{k=0}^{N_T-1} \Delta t |w_h^{k+1}|_1^2 = \sum_{k=0}^{N_T-1} \Delta t \sum_{x_i \in J_h} \sum_{x_j \in \Lambda_i} |Q_{ij}| (w_i^{k+1} - w_j^{k+1})^2 \leq C(T, w^0).$$

Proof. We use $\chi = \Delta t w_h^{k+1}$ in equation (3.2), sum from $k = 0 \rightarrow N_T - 1$, use the boundedness of $\|n_h^{k+1}\|_\infty$ and $\|w_h^{k+1}\|_\infty$ given by the Propositions 3.2 and 3.3, as well as the boundedness of ψ'_- to obtain the result. \square

Proposition 3.6 (Energy estimate on n_h^{k+1} and dissipation). *System (3.1) and (3.2) admits the a priori estimate*

$$\frac{1}{2} \|n_h^{k+1}\|_0^2 + \sum_{k=0}^{N_T-1} \Delta t \sum_{x_i \in J_h} \sum_{x_j \in \Lambda_i} \left(C_2 \hat{B}_{ij}^{k+1} + (\widehat{B\psi''})_{ij}^{k+1} \right) |Q_{ij}| (n_i^{k+1} - n_j^{k+1})^2 \leq C(T, w_h^0, n_h^0), \quad (3.15)$$

where C_2 is a small positive finite constant.

Proof. Multiplying equation (3.9) by n_i^{k+1} and summing over the nodes $x_i \in J_h$ leads to

$$\begin{aligned} & (n_h^{k+1} - n_h^k, n_h^{k+1})^h + \Delta t \sum_{x_i \in J_h} \sum_{x_j \in \Lambda_i} \left(\frac{\gamma}{\sigma} \hat{B}_{ij}^{k+1} + (\widehat{B\psi''})_{ij}^{k+1} \right) |Q_{ij}| (n_i^{k+1} - n_j^{k+1})^2 \\ &= \frac{\gamma}{\sigma} \Delta t \sum_{x_i \in J_h} \sum_{x_j \in \Lambda_i} \tilde{B}_{ij}^{k+1} |Q_{ij}| (w_i^{k+1} - w_j^{k+1})(n_i^{k+1} - n_j^{k+1}). \end{aligned}$$

The term on the right hand side can be bounded using Young's Inequality, and the boundedness of the mobility to obtain, for any $\kappa > 0$,

$$\begin{aligned} & \Delta t \sum_{x_i \in J_h} \sum_{x_j \in \Lambda_i} \frac{\gamma}{\sigma} \tilde{B}_{ij}^{k+1} |Q_{ij}| (w_i^{k+1} - w_j^{k+1})(n_i^{k+1} - n_j^{k+1}) \\ & \leq \Delta t \sum_{x_i \in J_h} \sum_{x_j \in \Lambda_i} \frac{\kappa \gamma^2}{2\sigma^2} |Q_{ij}| (w_i^{k+1} - w_j^{k+1})^2 + \frac{1}{2\kappa} \tilde{B}_{ij}^{k+1} |Q_{ij}| (n_i^{k+1} - n_j^{k+1})^2. \end{aligned}$$

Therefore, it leads to

$$\begin{aligned} & (n_h^{k+1} - n_h^k, n_h^{k+1})^h + \Delta t \sum_{x_i \in J_h} \sum_{x_j \in \Lambda_i} \left(\frac{\gamma}{\sigma} \hat{B}_{ij}^{k+1} - \frac{1}{2\kappa} \tilde{B}_{ij}^{k+1} + (\widehat{B\psi''})_{ij}^{k+1} \right) |Q_{ij}| (n_i^{k+1} - n_j^{k+1})^2 \\ & \leq \Delta t \sum_{x_i \in J_h} \sum_{x_j \in \Lambda_i} \frac{\kappa \gamma^2}{2\sigma^2} |Q_{ij}| (w_i^{k+1} - w_j^{k+1})^2. \end{aligned}$$

Therefore, from the boundedness of both \hat{B}_{ij}^{k+1} and \tilde{B}_{ij}^{k+1} , as well as the fact that $\hat{B}_{ij}^{k+1} \geq \tilde{B}_{ij}^{k+1}$ by definition, and since κ can be chosen arbitrarily, we know that it exists a finite constant $C_2 > 0$ such that $0 < C_2 \hat{B}_{ij}^{k+1} \leq \frac{\gamma}{\sigma} \hat{B}_{ij}^{k+1} - \frac{1}{2\kappa} \tilde{B}_{ij}^{k+1}$ (we recall that the term $\frac{\gamma}{\sigma}$ does not induce any difficulty since $\sigma < \gamma$). Using the property $(a - b)a \geq \frac{1}{2}(a^2 - b^2)$, Proposition 3.5 as well as the property (2.2), and summing for $k = 0 \rightarrow N_T - 1$, we obtain the result. \square

Corollary 3.7. *The following inequalities hold*

$$\sum_{k=0}^{N_T-1} \Delta t |\eta_h^{k+1}|_1^2 \leq C(T, n^0, w^0), \quad \text{and} \quad \sum_{k=0}^{N_T-1} \Delta t |\zeta_h^{k+1}|_1^2 \leq C(T, n^0, w^0). \quad (3.16)$$

Proof. From Proposition 3.6 and the non-negativity of $(\widehat{B\psi''})_{ij}^{k+1}$, we know that

$$\sum_{k=0}^{N_T-1} \Delta t \sum_{x_i \in J_h} \sum_{x_j \in \Lambda_i} \hat{B}_{ij}^{k+1} |Q_{ij}| (n_i^{k+1} - n_j^{k+1})^2 \leq C.$$

Furthermore, from the definition (3.12) and the definition of η provided by (1.15), we have

$$\begin{aligned} \sum_{k=0}^{N_T-1} \Delta t \sum_{x_i \in J_h} \sum_{x_j \in \Lambda_i} \hat{B}_{ij}^{k+1} |Q_{ij}| (n_i^{k+1} - n_j^{k+1})^2 &\geq \sum_{k=0}^{N_T-1} \Delta t \sum_{x_i \in J_h} \sum_{x_j \in \Lambda_i} |Q_{ij}| (\eta_i^{k+1} - \eta_j^{k+1})^2 \\ &= \sum_{k=0}^{N_T-1} \Delta t |\eta_h^{k+1}|_1^2. \end{aligned}$$

The same is obtained for $\sum_{k=0}^{N_T-1} \Delta t |\zeta_h^{k+1}|_1^2$ using similar arguments. \square

3.4. Existence of discrete solution

We prove well-posedness of our problem.

Theorem 3.8 (Well-posedness of the problem). *Let $d \leq 3$, and the spatio-temporal mesh satisfies the assumptions of Section 2. Then, System (3.1) and (3.2) with an initial condition satisfying (3.3) and (3.4), has a solution $\{n_h^{k+1}, w_h^{k+1}\} \in K^h \times V^h$ with*

$$0 \leq n_h^{k+1} \leq 1, \quad \text{in } \Omega. \quad (3.17)$$

Proof. The proof of the existence of a solution for the discrete problem relies on the use of Brouwer's fixed point theorem. A necessary step before applying the theorem is to change the unknown of our problem, we define $m_h^{k+1} = n_h^{k+1} - \alpha$, where $\alpha = \frac{1}{|\Omega|} \int_{\Omega} n^0 \, dx$. Therefore, the discrete problem (3.1) and (3.2) is equivalent to

$$F(\underline{m}^{k+1}) = \underline{m}^{k+1},$$

where $F : S \rightarrow S$ is an application defined on the space S with

$$S = \{\underline{m} \mid M_l \underline{m} \cdot (1, \dots, 1) = 0 \quad \text{and} \quad -\alpha \leq \underline{m} \leq 1 - \alpha\}.$$

The first constraint in the definition of the space S reflects the conservation of the initial mass while the second comes from Inequality (3.17). As a result, S is a convex and compact subspace of \mathbb{R}^{N_h} . Using the matrix formulation (3.5) and (3.6) for the problem P , the application F is defined by

$$F(\underline{m}^{k+1}) = \Delta t M_l^{-1} \left[U^{k+1} (\sigma Q + M_l)^{-1} \left(M_l \underline{n}^k - \frac{\gamma}{\sigma} M_l \underline{\psi}'_- \right) \right] - \Delta t R^{k+1} \underline{m}^{k+1} + \underline{n}^k - \alpha.$$

In the previous definition of F , we precise that the matrix U^{k+1} associated to the nonlinear convection term is given by

$$U_{ij}^{k+1} = \frac{\gamma}{\sigma} \int_{\Omega} \tilde{B}(m_h^{k+1} + \alpha) \nabla \chi_i \cdot \nabla \chi_j \, dx,$$

and for the diffusion term

$$R_{ij}^{k+1} = \int_{\Omega} \left(\frac{\gamma}{\sigma} \hat{B}(m_h^{k+1} + \alpha) + \widehat{(B\psi_+'')}(m_h^{k+1} + \alpha) \right) \nabla \chi_i \cdot \nabla \chi_j \, dx.$$

To check if F is continuous, we compute for $\underline{m}_1, \underline{m}_2 \in S$, $\|F(\underline{m}_1) - F(\underline{m}_2)\|$, and obtain

$$\begin{aligned} \|F(\underline{m}_1) - F(\underline{m}_2)\| &\leq \Delta t \left\| M_l^{-1} \left[(U(m_{h,1} + \alpha) - U(m_{h,2} + \alpha)) (\sigma Q + M_l)^{-1} \left(M_l \underline{n}^k - \frac{\gamma}{\sigma} M_l \psi_-' \right) \right] \right\| \\ &\quad + \Delta t \|R(m_{h,1} + \alpha) \underline{m}_1 - R(m_{h,2} + \alpha) \underline{m}_2\|. \end{aligned}$$

For the first term on the right-hand side of the previous inequality, we have

$$\|M_l^{-1}\| \leq C, \quad \|(\sigma Q + M_l)^{-1}\| \leq C, \quad \|M_l \underline{n}^k - \frac{\gamma}{\sigma} M_l \psi_-'\| \leq C,$$

from the properties of the mesh, the upper bound presented by Varah [36] for the inverse of M-matrices, and our assumptions on n_h^k and w_h^k as well as the function $\psi_-(\cdot)$. Lastly, from the definitions of the matrix $U(\cdot)$ and of the convective flux G (see Def. 3.1), we have

$$\|U(m_{h,1} + \alpha) - U(m_{h,2} + \alpha)\| \leq C \|\underline{m}_1 - \underline{m}_2\|.$$

The same applies using the properties of the matrix $R(\cdot)$ to obtain

$$\|R(m_{h,1} + \alpha) \underline{m}_1 - R(m_{h,2} + \alpha) \underline{m}_2\| \leq C \|\underline{m}_1 - \underline{m}_2\|.$$

Altogether, using the properties of the standard finite element matrices and the fact that the mobility $\tilde{B}(m_{h,\varepsilon}^k + \alpha)$ is bounded, we obtain

$$\|F(\underline{m}_1) - F(\underline{m}_2)\| \leq C(\Delta t, h) \|\underline{m}_1 - \underline{m}_2\|,$$

which proves that the mapping F is Lipschitz continuous. Therefore, applying Brouwer's fixed point theorem, we know that it exists a solution $\underline{m}^{k+1} \in S$ such that $F(\underline{m}^{k+1}) = \underline{m}^{k+1}$. Therefore, it exists a $m_h^{k+1} \in V^h$ that gives the existence of a pair $\{n_h^{k+1}, w_h^{k+1}\} \in V^h \times V^h$ solution of the Problem P . \square

3.5. Compactness estimates

We now derive estimates for the time and space translates of the discrete solutions. This results follow the lines of the works [14, 15]. We first define the space and time discrete spaces $V_{\Delta t}^h$ and $\hat{V}_{\Delta t}^h$. These sets are composed of piecewise constant functions in time with values in V^h and \hat{V}^h respectively. To study the convergence as $h, \Delta t \rightarrow 0$, we use an index m such that, as $m \rightarrow \infty$, $h_m, \Delta t_m \rightarrow 0$. Therefore, we study the convergence of sequences of functions in the spaces $V_{\Delta t_m}^{h_m}$ and $\hat{V}_{\Delta t_m}^{h_m}$.

We define by $\eta_{h_m, \Delta t_m}$, $\zeta_{h_m, \Delta t_m}$ and $w_{h_m, \Delta t_m}$ the piecewise affine in space and piecewise constant in time approximation of the functions η , ζ and w . We also denote by $\hat{\eta}_{h_m, \Delta t_m}$, $\hat{\zeta}_{h_m, \Delta t_m}$ and $\hat{w}_{h_m, \Delta t_m}$, the piecewise constant in space and time corresponding approximations. For each node $x_i \in J_h$ and time $t = t^k$, we have the notation $\eta_{h_m, \Delta t_m}(x_i, t^k) = \hat{\eta}_{h_m, \Delta t_m}(x_i, t^k) = \eta_i^k$. The associated vector containing all the value for all nodes is denoted by $\underline{\eta}^k$.

Time translate estimates. We start by estimates on the time translates. We denote by $Q_{T-\tau} = \Omega \times (0, T-\tau)$, for all $\tau \in (0, T)$. We have the following result

Lemma 3.9 (Time translate estimates). *There are constants $C_{\eta,t}$, $C_{\zeta,t}$ and $C_{w,t}$ independent of h and τ such that the following inequalities hold*

$$\int_{Q_{t-\tau}} |\hat{\eta}_{h_m, \Delta t_m}(x, t+\tau) - \hat{\eta}_{h_m, \Delta t_m}(x, t)|^2 dx dt \leq C_{\eta,t}(\tau + \Delta t), \quad (3.18)$$

$$\int_{Q_{t-\tau}} \left| \hat{\zeta}_{h_m, \Delta t_m}(x, t+\tau) - \hat{\zeta}_{h_m, \Delta t_m}(x, t) \right|^2 dx dt \leq C_{\zeta,t}(\tau + \Delta t), \quad (3.19)$$

$$\int_{Q_{t-\tau}} |\hat{w}_{h_m, \Delta t_m}(x, t+\tau) - \hat{w}_{h_m, \Delta t_m}(x, t)|^2 dx dt \leq C_{w,t}(\tau + \Delta t). \quad (3.20)$$

Proof. The proof is similar to the proof of Lemma 4.3 in [14]. We give here the details for the sake of clarity since the equation under study is different.

We start by defining the quantity

$$A_m(t) = \int_{\Omega} |\hat{\eta}_{h_m, \Delta t_m}(x, t+\tau) - \hat{\eta}_{h_m, \Delta t_m}(x, t)|^2 dx, \quad \forall t \in (0, T-\tau).$$

Since the functions $\hat{\eta}_{h_m, \Delta t_m}$ are piecewise constant in time, we define $v(t)$, for $t \in (0, T]$, the unique integer such that $t^{v(t)} < t \leq t^{v(t)+1}$. Therefore, we write $\forall t \in (0, T-\tau)$,

$$\begin{aligned} A_m(t) &= \left(\underline{\eta}^{v(t+\tau)+1} - \underline{\eta}^{v(t)+1} \right)^T M_l \left(\underline{\eta}^{v(t+\tau)+1} - \underline{\eta}^{v(t)+1} \right), \\ &= \sum_{i=1}^{N_h} \left(\eta_i^{v(t+\tau)+1} - \eta_i^{v(t)+1} \right)^2 M_{l,ii}. \end{aligned}$$

However, from definition (1.15), we know that it exists a constant C such that

$$\left(\eta_i^{v(t+\tau)+1} - \eta_i^{v(t)+1} \right)^2 \leq C \left(n_i^{v(t+\tau)+1} - n_i^{v(t)+1} \right) \left(\eta_i^{v(t+\tau)+1} - \eta_i^{v(t)+1} \right).$$

Then, using the definition of the integer $v(\cdot)$, we find

$$C \left(n_i^{v(t+\tau)+1} - n_i^{v(t)+1} \right) \left(\eta_i^{v(t+\tau)+1} - \eta_i^{v(t)+1} \right) = C \sum_{k=v(t)+1}^{v(t+\tau)} (n_i^{k+1} - n_i^k) \left(\eta_i^{v(t+\tau)+1} - \eta_i^{v(t)+1} \right).$$

From the first equation of the scheme (3.1), we obtain

$$\begin{aligned} A_m(t) &\leq -C \sum_{k=v(t)+1}^{v(t+\tau)} \Delta t_m \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} \left(\frac{\gamma}{\sigma} \hat{B}_{ij}^{k+1} + \widehat{(B\psi''_+)}_{ij}^{k+1} \right) |Q_{ij}| (n_i^{k+1} - n_j^{k+1}) \\ &\quad \times \left(\left(\eta_i^{v(t+\tau)+1} - \eta_j^{v(t+\tau)+1} \right) - \left(\eta_i^{v(t)+1} - \eta_j^{v(t)+1} \right) \right) \\ &\quad + C \sum_{k=v(t)+1}^{v(t+\tau)} \frac{\gamma \Delta t_m}{\sigma} \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} \tilde{B}_{ij}^{k+1} |Q_{ij}| (w_i^{k+1} - w_j^{k+1}) \\ &\quad \times \left(\left(\eta_i^{v(t+\tau)+1} - \eta_j^{v(t+\tau)+1} \right) - \left(\eta_i^{v(t)+1} - \eta_j^{v(t)+1} \right) \right). \end{aligned}$$

Using Young's inequality, we have

$$A_m(t) \leq C (A_{1,m}(t) + A_{2,m}(t) + A_{3,m}(t) + A_{4,m}(t)),$$

where

$$\begin{aligned} A_{1,m}(t) &= \frac{1}{2} \sum_{k=v(t)+1}^{v(t+\tau)} \Delta t_m \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} \left(\frac{\gamma}{\sigma} \hat{B}_{ij}^{k+1} + (\widehat{B\psi''})_{ij}^{k+1} \right) |Q_{ij}| (n_i^{k+1} - n_j^{k+1})^2, \\ A_{2,m}(t) &= \left(\frac{\gamma}{\sigma} \|b\|_\infty + \frac{1}{2} \|b\psi''\|_\infty \right) \sum_{k=v(t)+1}^{v(t+\tau)} \Delta t_m \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} |Q_{ij}| \left(\eta_i^{v(t+\tau)+1} - \eta_j^{v(t+\tau)+1} \right)^2, \\ A_{3,m}(t) &= \left(\frac{\gamma}{\sigma} \|b\|_\infty + \frac{1}{2} \|b\psi''\|_\infty \right) \sum_{k=v(t)+1}^{v(t+\tau)} \Delta t_m \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} |Q_{ij}| \left(\eta_i^{v(t)+1} - \eta_j^{v(t)+1} \right)^2, \\ A_{4,m}(t) &= \frac{1}{2} \sum_{k=v(t)+1}^{v(t+\tau)} \Delta t_m \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} \frac{\gamma}{\sigma} \tilde{B}_{ij}^{k+1} |Q_{ij}| (w_i^{k+1} - w_j^{k+1})^2. \end{aligned}$$

To handle the first sum in each of these quantities, we introduce some additional notations. We define $\rho(k, t, \tau)$ the characteristic function such that

$$\rho(k, t, \tau) = \begin{cases} 1, & \text{if } t < k\Delta t_m \leq t + \tau, \\ 0, & \text{otherwise.} \end{cases}$$

Hence, we have

$$\int_0^{T-\tau} \rho(k, t, \tau) dt = \int_{t^k-\tau}^{t^k} dt = \tau, \quad \text{and} \quad \sum_{k=v(t)+1}^{v(t+\tau)} \Delta t_m = \sum_{k:t^k < t^k \leq t+\tau} t^{k+1} - t^k \leq \tau + \Delta t_m.$$

Furthermore, for any family of real non-negative values $(a^k)_{k \in \{0, \dots, N_T\}}$, we have

$$\int_0^{T-\tau} \sum_{k=v(t)+1}^{v(t+\tau)} \Delta t_m a^{k+1} dt = \int_0^{T-\tau} \sum_{k=0}^{N_T-1} \Delta t_m a^{k+1} \rho(k, t, \tau) dt = \tau \sum_{k=0}^{N_T-1} \Delta t_m a^{k+1}.$$

From Propositions 3.5 and 3.6, we find

$$\int_0^{T-\tau} A_{1,m}(t) dt \leq C(\tau + \Delta t_m), \quad \text{and} \quad \int_0^{T-\tau} A_{4,m}(t) dt \leq C\tau \leq C(\tau + \Delta t_m).$$

Then, using arguments similar to Proposition 9.3 in [25], we have for any family of real non-negative values $(a^k)_{k \in \{0, \dots, N_T\}}$,

$$\int_0^{T-\tau} \sum_{k=v(t)+1}^{v(t+\tau)} \Delta t_m a^{v(t+\tau)+1} dt \leq (\tau + \Delta t_m) \sum_{k=0}^{N_T-1} (\Delta t_m a^{k+1}).$$

Using the previous argument and Corollary 3.7, we obtain

$$A_{2,m} + A_{3,m} \leq C(\tau + \Delta t).$$

This achieves the proof of Inequality (3.18). The proofs of Inequalities (3.19) and (3.20) are very similar and we do not repeat them. \square

Remark 3.10. From the previous lemma, we can easily give the time translate estimates over \mathbb{R}^{d+1} . Indeed, extending by zero the functions $\hat{\eta}_{h_m, \Delta t_m}$, $\hat{\zeta}_{h_m, \Delta t_m}$ and $\hat{w}_{h_m, \Delta t_m}$ outside of $\Omega \times (0, T)$, we obtain

$$\begin{cases} \int_{\mathbb{R}^{d+1}} |\hat{\eta}_{h_m, \Delta t_m}(x, t + \tau) - \hat{\eta}_{h_m, \Delta t_m}(x, t)|^2 dx dt \leq C_{\eta, t}(\tau + \Delta t), \\ \int_{\mathbb{R}^{d+1}} |\hat{\zeta}_{h_m, \Delta t_m}(x, t + \tau) - \hat{\zeta}_{h_m, \Delta t_m}(x, t)|^2 dx dt \leq C_{\zeta, t}(\tau + \Delta t), \\ \int_{\mathbb{R}^{d+1}} |\hat{w}_{h_m, \Delta t_m}(x, t + \tau) - \hat{w}_{h_m, \Delta t_m}(x, t)|^2 dx dt \leq C_{w, t}(\tau + \Delta t). \end{cases}$$

To obtain this result, we use

$$\begin{aligned} \int_{\mathbb{R}^{d+1}} |\hat{\eta}_{h_m, \Delta t_m}(x, t + \tau) - \hat{\eta}_{h_m, \Delta t_m}(x, t)|^2 dx dt &= \int_{Q_{t-\tau}} |\hat{\eta}_{h_m, \Delta t_m}(x, t + \tau) - \hat{\eta}_{h_m, \Delta t_m}(x, t)|^2 dx dt \\ &\quad + \int_{T-\tau}^T \int_{\Omega} |\hat{\eta}_{h_m, \Delta t_m}(x, t)|^2 dx dt, \end{aligned}$$

with Lemma 3.9 and the L^∞ -bounds on $\hat{\eta}_{h_m, \Delta t_m}$, $\hat{\zeta}_{h_m, \Delta t_m}$, and $\hat{w}_{h_m, \Delta t_m}$.

Space translate estimates. We now turn to the space translate estimates.

Lemma 3.11 (Space translate estimates). *There are three constants $C_{\eta, s}$, $C_{\zeta, s}$ and $C_{w, s}$ independent of m and y such that*

$$\int_0^T \int_{\mathbb{R}^d} |\hat{\eta}_{h_m, \Delta t_m}(x + y, t) - \hat{\eta}_{h_m, \Delta t_m}(x, t)| dx dt \leq C_{\eta, s}(|y| + h_m), \quad (3.21)$$

$$\int_0^T \int_{\mathbb{R}^d} |\hat{\zeta}_{h_m, \Delta t_m}(x + y, t) - \hat{\zeta}_{h_m, \Delta t_m}(x, t)| dx dt \leq C_{\zeta, s}(|y| + h_m), \quad (3.22)$$

$$\int_0^T \int_{\mathbb{R}^d} |\hat{w}_{h_m, \Delta t_m}(x + y, t) - \hat{w}_{h_m, \Delta t_m}(x, t)| dx dt \leq C_{w, s}(|y| + h_m). \quad (3.23)$$

Proof. The proof of this result follows Lemmas 4.1 and 4.2 in [14]. For the sake of clarity, we here present the main steps.

First, using Corollary 3.7, we know that $\|\nabla \eta_{h_m, \Delta t_m}\|_{(L^2(\Omega \times (0, T)))^d}$ is bounded. Since the time and space domain considered is of finite measure, Hölder inequality indicates that $\|\nabla \eta_{h_m, \Delta t_m}\|_{(L^1(\Omega \times (0, T)))^d}$ is also bounded. Hence, since $\eta_{h_m, \Delta t_m}$ is bounded as well, its extension by zeros outside $\Omega \times (0, T)$ lies in $L^\infty \cap BV(\mathbb{R}^{d+1})$. Therefore, we have

$$\int_0^T \int_{\mathbb{R}^d} |\eta_{h_m, \Delta t_m}(t, x + y) - \eta_{h_m, \Delta t_m}(t, x)| dx dt \leq C|y|,$$

from which, using the triangular inequality and Remark 3.12 stated below, we find Inequality (3.21). The same is found for Inequality (3.22) and Inequality (3.23) following the same arguments. \square

Remark 3.12. One of the important tool that we did not present in the previous proof is

$$\int_{\Omega_T} |\eta_{h_m, \Delta t_m}(t, x + y) - \hat{\eta}_{h_m, \Delta t_m}(t, x + y)| dx dt \leq Ch_m, \quad (3.24)$$

found from the use of Lemma A.2 in [14].

3.6. Convergence analysis

This section is organized as follow, we first apply Fréchet-Kolmogorov theorem to show strong convergence in $L^1(\Omega_T)$, then we show that the limit is a solution of the continuous RDCH system.

Lemma 3.13 (Strong convergence in L^1). *As $m \rightarrow \infty$, we can extract subsequences of $(\hat{\eta}_{h_m, \Delta t_m})_{m \geq 1}$, $(\hat{\zeta}_{h_m, \Delta t_m})_{m \geq 1}$ and $(\hat{w}_{h_m, \Delta t_m})_{m \geq 1}$ converging in $L^1(\Omega_T)$ to the limit functions $\eta(n) \in L^2(0, T; H^1(\Omega))$, $\zeta(n) \in L^2(0, T; H^1(\Omega))$ and $w \in L^2(0, T; H^1(\Omega))$ such that*

$$\hat{\eta}_{h_m, \Delta t_m}, \hat{\zeta}_{h_m, \Delta t_m}, \hat{w}_{h_m, \Delta t_m} \rightarrow \eta(n), \zeta(n), w, \quad \text{strongly in } L^1(\Omega_T), \quad (3.25)$$

$$\eta_{h_m, \Delta t_m}, \zeta_{h_m, \Delta t_m}, w_{h_m, \Delta t_m} \rightharpoonup \eta(n), \zeta(n), w, \quad \text{weakly in } L^2(0, T; H^1(\Omega)). \quad (3.26)$$

Furthermore, we have

$$\hat{n}_{h_m, \Delta t_m} \rightarrow n, \quad \text{a.e. in } \Omega_T, \quad \text{and strongly in } L^p(\Omega_T), \quad \text{for } p < +\infty. \quad (3.27)$$

Proof. Again, the proof is similar to the proof of Lemma 4.5 in [14]. Indeed, from Lemmas 3.9 and 3.11, and the boundedness provided by Corollary 3.7, we know that the sequence $(\hat{\eta}_{h_m, \Delta t_m})_{m \geq 0}$ satisfies the necessary assumptions to use Fréchet-Kolmogorov theorem. Therefore, we know that $(\hat{\eta}_{h_m, \Delta t_m})_{m \geq 0}$ is relatively compact in $L^1(\Omega_T)$, which implies (3.25) (the result for the sequences $(\hat{\zeta}_{h_m, \Delta t_m})_{m \geq 0}$ and $(\hat{w}_{h_m, \Delta t_m})_{m \geq 0}$ follows similar arguments). Furthermore, the limit is in $L^2(0, T; H^1(\Omega))$ since from the use of Corollary 3.7, we know that $(\eta_{h_m, \Delta t_m})_{m \geq 0}$ also converges weakly in $L^2(0, T; H^1(\Omega))$ to a limit. Since we know from Inequality (3.24), that the sequences $(\eta_{h_m, \Delta t_m})_{m \geq 0}$ and $(\hat{\eta}_{h_m, \Delta t_m})_{m \geq 0}$ have the same limit and $(\hat{\eta}_{h_m, \Delta t_m})_{m \geq 0}$ already converges to a limit in $L^1(\Omega_T)$, uniqueness of the limit gives the result. The same applies for the quantities $\hat{\zeta}_{h_m, \Delta t_m}$, $\hat{w}_{h_m, \Delta t_m}$ and $\zeta(n)_{h_m, \Delta t_m}$, $w_{h_m, \Delta t_m}$ such that we obtain the convergences (3.25) and (3.26).

To prove the convergence (3.27), we start by stating that the inverse η^{-1} of the continuous function η exists and is continuous. Furthermore, we have that $\hat{n}_{h_m, \Delta t_m}$ is bounded in L^∞ from Proposition 3.2. Therefore, applying the dominated convergence theorem to $\hat{n}_{h_m, \Delta t_m} = \eta^{-1}(\hat{\eta}_{h_m, \Delta t_m})$, we arrive to the convergence (3.27) in which the limit is defined as the unique function $n(t, x) = \eta^{-1}(\eta)$ (η being the limit function in (3.25)). \square

Theorem 3.14 (Limit system). *The limit of $(\hat{n}_{h_m, \Delta t_m}, \hat{w}_{h_m, \Delta t_m}, \hat{\eta}_{h_m, \Delta t_m}, \hat{\zeta}_{h_m, \Delta t_m})$ denoted (n, w, η, ζ) is solution of the RDCH system in the sense of Definition 1.1.*

Proof. The proof of this theorem is very close to Section 5 in [15]. For the sake of clarity, our proof can be found in Appendix A. \square

4. LINEARIZED SEMI-IMPLICIT NUMERICAL SCHEME

To restrain the computational time of the simulation of the RDCH model within reasonable bounds, we propose a linearized semi-implicit version of the numerical scheme. We apply the CVFE framework to the original formulation of the RDCH problem this time. The problem now reads:

For each $k = 0, \dots, N_T - 1$, find $\{n_h^{k+1}, \varphi_h^{k+1}\}$ in $K^h \times V^h$ such that

$$\left(\frac{n_h^{k+1} - n_h^k}{\Delta t}, \chi_1 \right)^h + (b(n_h^k) \psi_+''(n_h^k) \nabla n_h^{k+1}, \nabla \chi_1) = - \left(\tilde{B}(n_h^k) \nabla \varphi_h^{k+1}, \nabla \chi_1 \right), \quad \forall \chi_1 \in V^h, \quad (4.1a)$$

$$\sigma(\nabla \varphi_h^{k+1}, \nabla \chi_2) + (\varphi_h^{k+1}, \chi_2)^h = \gamma(\nabla n_h^k, \nabla \chi_2) + \left(\psi_-'(n_h^k - \frac{\sigma}{\gamma} \varphi_h^k), \chi_2 \right)^h, \quad \forall \chi_2 \in V^h. \quad (4.1b)$$

We define the following finite elements matrices

$$U_{ij}^k = \int_{\Omega} \tilde{B}_{ij}^k \nabla \chi_i \cdot \nabla \chi_j \, dx, \quad \text{for } i, j = 1, \dots, N_h, \quad (4.2)$$

and

$$L_{ij}^k = \int_{\Omega} b(n_{h,\epsilon}^k) \psi''_+(n_h^k) \nabla \chi_i \cdot \nabla \chi_j \, dx, \quad \text{for } i, j = 1, \dots, N_h. \quad (4.3)$$

We write the matrix form of equation (4.1a)

$$(M_l + \Delta t L^k) \underline{n}^{k+1} = -\Delta t U^k \underline{\varphi}^{k+1} + M_l \underline{n}^k,$$

and since U^k has zero row sum, we can rewrite the previous equation for each node $x_i \in J_h$,

$$M_{l,ii} n_i^{k+1} = M_{l,ii} n_i^k - \Delta t \sum_{x_j \in \Lambda_i} [L_{ij}^k (n_j^{k+1} - n_i^{k+1}) + U_{ij}^k (\varphi_j^{k+1} - \varphi_i^{k+1})], \quad (4.4)$$

where Λ_i is the set of nodes connected to the node x_i by an edge. In the definition of (4.2) we compute the mobility coefficient in function of the direction of $\nabla \varphi_h^{k+1}$. As for the nonlinear case, the mobility coefficient is given by

$$\tilde{B}_{ij}^k = \begin{cases} n_i^k (1 - n_j^k)^2, & \text{if } \varphi_i^{k+1} - \varphi_j^{k+1} > 0, \\ n_j^k (1 - n_i^k)^2, & \text{otherwise.} \end{cases}$$

Even though we cannot redo the same analysis as for the nonlinear scheme, we can establish the existence and the nonnegativity of discrete solutions of (4.1a) and (4.1b).

Theorem 4.1 (Well-posedness of linear upwind scheme). *Let $\Omega \subset \mathbb{R}^d$, $d = 1, 2, 3$, and assume that \mathcal{T}^h is a quasi-uniform acute mesh of Ω , and the condition*

$$\frac{(d+1) G_h \Delta t}{\kappa_h^2} \max_{\substack{x_i \in J_h \\ x_j \in \Lambda_i}} (\varphi_j^{k+1} - \varphi_i^{k+1}) < 1, \quad (4.5)$$

(where Λ_i is the set of nodes connected to the node x_i by an edge) is satisfied. Then, the linear finite element scheme (4.1a) and (4.1b) with initial condition $n_h^0 \in K^h$ admits a unique solution $\{n_h^{k+1}, \varphi_h^{k+1}\} \in K^h \times V^h$ satisfying

$$0 \leq n_h^{k+1} < 1.$$

Proof. Step 1. Existence of a unique solution. Assuming that $\{n_h^k, \varphi_h^k\} \in K^h \times V^h$, from the Lax-Milgram theorem, it exists a unique solution $\varphi_h^{k+1} \in V^h$ of equation (4.1b) and equation (4.1a) admits a unique solution $n_h^{k+1} \in V^h$. Therefore, it exists a unique pair of discrete solutions $\{n_h^{k+1}, \varphi_h^{k+1}\} \in V^h \times V^h$ for the system (4.1a) and (4.1b). Next, we need to prove that n_h^{k+1} is nonnegative and bounded from above by 1.

Step 2. Nonnegativity and upper bound on n_h^{k+1} for $d = 1, 2, 3$. First, from the fact that $(M_l + \Delta t L^k)$ is a M-matrix, we know that its inverse is non-negative, i.e. $(M_l + \Delta t L^k)^{-1} \geq 0$. Therefore, to preserve the non-negativity of n_h^{k+1} , we need that

$$M_l \underline{n}^k - \Delta t U^k \underline{\varphi}^{k+1} \geq 0.$$

For every node $x_i \in J_h$, the previous condition reads

$$|D_i| n_i^k - \Delta t \sum_{x_j \in \Lambda_i} B_{ij}^k Q_{ij} (\varphi_i^{k+1} - \varphi_j^{k+1}) \geq 0,$$

where Λ_i is the set of nodes connected to the node x_i by an edge. From the fact that the mesh is acute, we know that Q_{ij} is negative. Therefore, using the definition of the mobility coefficient (3.11), we need to focus on the case $\varphi_j^{k+1} - \varphi_i^{k+1} < 0$. In that situation, we have

$$n_i^k - \frac{\Delta t}{|D_i|} \sum_{x_j \in \Lambda_i} n_i^k (1 - n_j^k)^2 Q_{ij} (\varphi_j^{k+1} - \varphi_i^{k+1}) \geq 0.$$

However, from (2.7), we find the following condition to ensure the non-negativity of n_h^{k+1}

$$\frac{(d+1) G_h \Delta t}{\kappa_h^2} \max_{\substack{x_i \in J_h \\ x_j \in \Lambda_i}} (\varphi_j^{k+1} - \varphi_i^{k+1}) \leq 1.$$

Then, we need to prove that for every node $x_i \in J_h$ we have $n_i^{k+1} < 1$. We use the upper bound presented by Varah [36] for the inverse of M-matrices, and write

$$\left\| \left(\frac{M_l}{\Delta t} + L^k \right)^{-1} \right\|_{\infty} \leq \frac{\Delta t}{M_{l,ii}}.$$

Therefore, to retrieve the upper bound on the discrete solution, the condition

$$n_i^k - \frac{\Delta t}{|D_i|} \sum_{x_j \in \Lambda_i} n_j^k (1 - n_i^k)^2 Q_{ij} (\varphi_j^{k+1} - \varphi_i^{k+1}) < 1,$$

has to be satisfied. Note in the previous equation that we have considered the case $\varphi_j^{k+1} - \varphi_i^{k+1} > 0$ since in the other case the bound will be satisfied trivially. Then, subtracting n_i^k from both sides of the previous inequality, we obtain

$$-\frac{\Delta t}{|D_i|} \sum_{x_j \in \Lambda_i} n_j^k (1 - n_i^k) Q_{ij} (\varphi_j^k - \varphi_i^k) < 1,$$

and we retrieve the same condition as before with a strict inequality.

Altogether, we proved the existence of a unique solution $\{n_h^{k+1}, \varphi_h^k\} \in K^h \times V^h$ for the system (4.1a) and (4.1b) with $0 \leq n_h^{k+1} < 1$ if the stability condition (4.5) is satisfied. \square

5. NUMERICAL SIMULATIONS

In this section, we use the previously presented linear scheme (4.1a) and (4.1b) for the RDCH system. The numerical simulations are performed using the MATLAB software. At each time step the matrices U^k and L^k are reassembled. The linear system is solved using the function *linsolve* of the MATLAB software. This function uses the LU factorization.

Even though we are presenting numerical results obtained using the linear scheme (4.1a) and (4.1b), the evolution of the energy during the simulations is given from the computation of the discrete formulation of the continuous energy

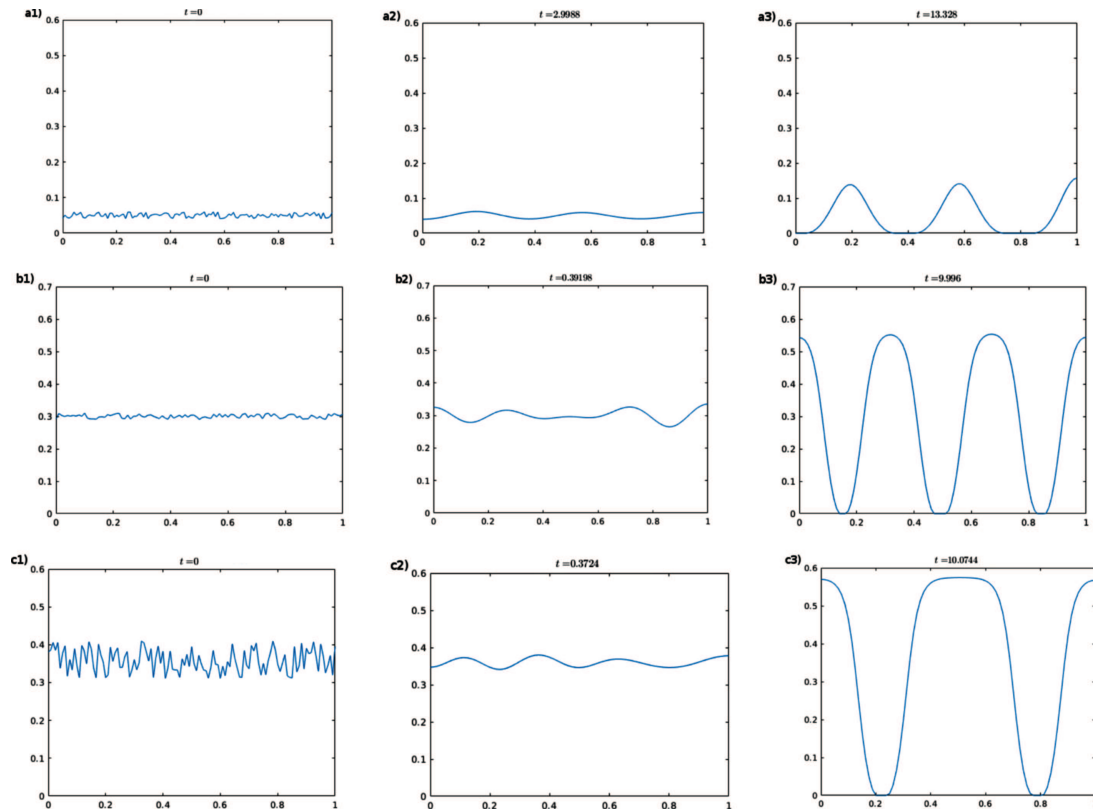
$$E(n_h^{k+1}, \varphi_h^{k+1}) := \int_{\Omega} \frac{\gamma}{2} \left| \nabla \left(n_h^{k+1} - \frac{\sigma}{\gamma} \varphi_h^{k+1} \right) \right|^2 + \frac{\sigma}{2\gamma} |\varphi_h^{k+1}|^2 + \psi_+(n_h^{k+1}) + \psi_- \left(n_h^{k+1} - \frac{\sigma}{\gamma} \varphi_h^{k+1} \right) dx.$$

First of all, we present test cases in one and two dimensions to validate our method. The physical properties of the solutions such as the shape of the aggregates, the energy decay, the mass preservation and the non-negativity of the solution are the key characteristics we need to observe to validate our method. A comparison with previous results from the literature is also of main importance. The reference used for this study is the work of Agosti *et al.* [2]. The analysis of the long-time behavior of the solutions of the RDCH equation [32] gives us some insights about what we should observe at the end of the simulations. The solutions should evolve to steady-states that are minimizers of the energy functional. Depending on the initial mass, three regions of the cell density should appear. The first being the region of absence of cells, the second the continuous interface linking the bottom and the top of the aggregates. If the initial mass is sufficiently large, the third expected region is a plateau of the cell density close to $n = n^*$. This evolution is related to the clustering of tumor cells in *in-vitro* biological experiments (see *e.g.* [3, 10]).

The study of the effect of the regularization on the numerical scheme is the purpose of the last subsection.

TABLE 1. Parameters of the 1D test case.

Parameters	
γ	$(0.014)^2$
Δt	0.1γ
h	0.01
n^0	$\{0.05, 0.3, 0.36\}$
n^*	0.6
σ	5.10^{-5}

FIGURE 3. Solution n_h at 3 different times with $n^0 = 0.05$ (a1, a2, a3), $n^0 = 0.3$ (b1, b2, b3) and $n^0 = 0.36$ (c1, c2, c3).

5.1. Numerical results: Test cases

1D test cases

Table 1 summarizes the parameters used for the one dimensional test cases. The initial cell density is a uniform distributed random perturbation around the values n^0 . Figure 3 show the evolution in time of the solutions n_h for the three different initial masses.

We can observe that the solution for each of the three test cases remains nonnegative and the mass is conserved throughout the simulations. From Figure 4, we observe that the energies decrease monotonically for the three

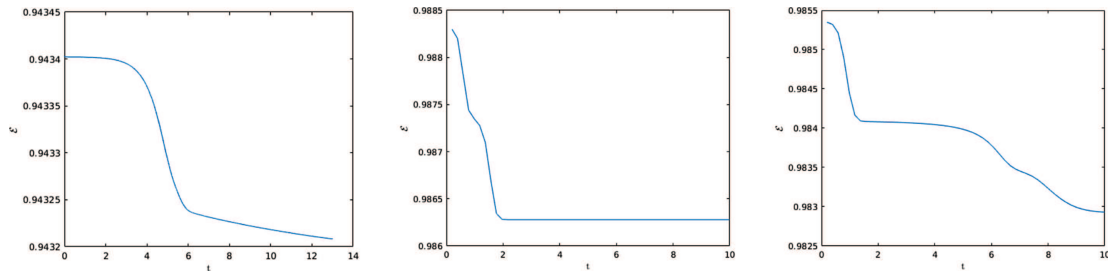


FIGURE 4. Evolution of discrete energy through time for the 3 initial conditions (from *left to right* $n^0 = \{0.05, 0.3, 0.36\}$).

simulations but at different speeds. They all display at the end of the computation a stable (or metastable) state that is a global (or respectively a local) minimizer of the discrete energy.

For the initial condition $n^0 = 0.3$ (Figs. 3(b1, b2, b3) and Fig. 4 at the middle), the energy decreases rapidly and reaches a plateau showing that the solution evolves rapidly to a steady state. The solution at $t = 10$ presents aggregates that are not saturated (*i.e.* the maximum density is below n^*). The explanation behind this observation is that the initial mass is not sufficiently large for the system to produce any saturated aggregates. However, the clusters appear to be of similar thickness and are relatively symmetrically distributed in the domain.

For $n^0 = 0.36 > n^*/2$ (Figs. 3(c1, c2, c3)), the aggregates are thicker. The top of the aggregate located at the center of the domain is flatter than for the other simulation. The maximum density is closer to the value n^* than for the initial condition $n^0 = 0.3$. Likewise, the symmetry in the domain is respected. Using Figure 4 on the right, we observe that at different times, the energy evolves through several meta-stable equilibria. This reflects the fact that the solution went to different meta-stable states before reaching a stable equilibrium that better minimizes the energy.

For the initial condition $n^0 = 0.05$ (Figs. 3(a1, a2, a3)), the shape of the final solution is different. The aggregates appear to be thinner and far from each other. The symmetry is not retrieved in the domain. Furthermore, from Figure 4 (on the left), we can observe that the evolution of the solution is slow compared to the other initial conditions. The energy seems to be constant in the first moment of the simulation (*i.e.* after the spinodal decomposition phase). The slow evolution of the solution is explained from the fact that the mobility is degenerate and the amount of mass available in the domain is small. Using Figure 4, we can also see that the energy continues to decrease even at the end of the simulation. To keep comparable simulation times, we did not reach the complete steady state.

Let us compare qualitatively these results with the ones obtained in [2] for the one dimensional case. For the two test cases $n^0 = 0.3$ and $n^0 = 0.36$, there is no differences in the shape the aggregates or the distribution of the mass in the domain. For $n^0 = 0.05$, some small discrepancies with the final solutions are observed. In particular, the symmetry of the aggregates in the domain is not respected in our case whereas it is in the reference work. We must stress that doing other simulations, the symmetry was sometimes reached at the time $t \approx 100$ for the initial condition $n^0 = 0.05$. The reason is that the system will evolve to respect the symmetry but the time at which this stable-steady state is reached depends on the initial distribution of the cell density.

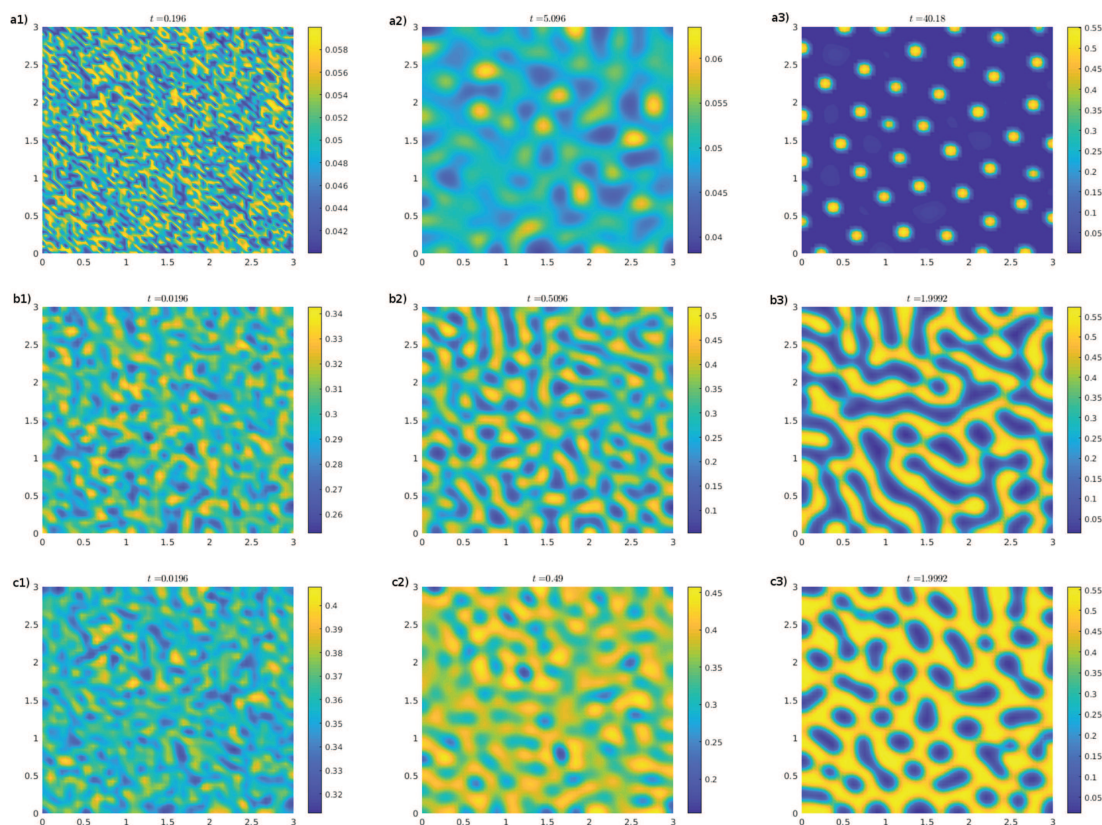
Altogether, the solutions obtained at the end the three simulations are in accordance with the description of the steady-states made in [32]. The three regions of interest are indeed retrieved at the end of each simulation.

2D test cases

For the two-dimensional test cases, the domain is a square of length $L = 1$. The initial density is computed in the same way as for the one-dimensional test cases *i.e.* a random uniformly distributed perturbation around n^0 . The summary of the values of parameters can be found in Table 2. Figure 5 depicts the results of the three

TABLE 2. Parameters of the test cases.

Parameters	
γ	0.014^2
Δt	2γ
h	$1/64$
n^0	$[0.05, 0.3, 0.36]$
n^*	0.6
σ	10^{-5}

FIGURE 5. Solution n_h at 3 different times with $n^0 = 0.05$ (a1, a2, a3), $n^0 = 0.3$ (b1, b2, b3) and $n^0 = 0.36$ (c1, c2, c3).

test cases with different initial masses. The three simulations respect the nonnegativity of the cell density, the conservation of the initial mass and the monotonic decay of the discrete energy. However, different shapes can be observed for the aggregates.

Figures 5(a1, a2, a3) show the evolution of the solution through time for the small initial mass $n^0 = 0.05$. Starting from a uniform random distribution of the cell density in the domain, the solution evolves into a more organized configuration. Progressively, a separation of the two phases of the mixture occurs. At the end of the simulation, small clusters are formed. They display a circular shape and are of similar width. The organization of the clusters in the domain tries to maximize the distance between each other. Using the Figure 6 (left),

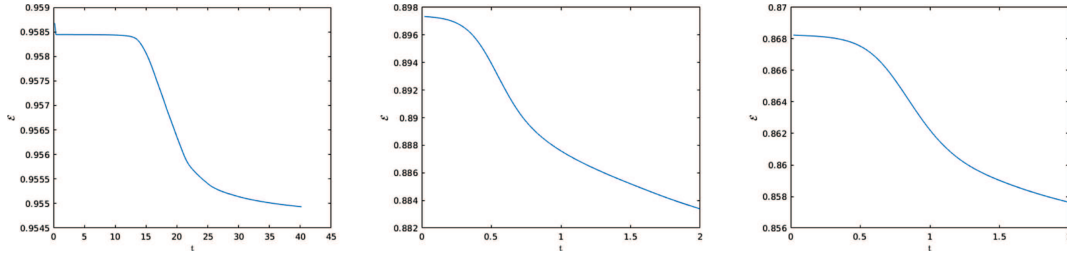


FIGURE 6. Evolution of discrete energy through time for the 3 initial conditions (from *left* to *right* $n^0 = \{0.05, 0.3, 0.36\}$).

we observe a drop of the energy in the first moments of the simulation denoting a fast reorganization of the randomly distributed initial condition. Then, the solution appears to evolve very slowly, *i.e.* a meta-stable state was reached. A second drop of the energy follows around $t \approx 15$, the system enters the “coarsening” phase: the small aggregates become more dense and merge with others. At the end, the evolution is very slow. The system continues to rearrange but due to the degeneracy of the mobility and the small amount of initial mass this process is very slow.

Figures 5(b1, b2, b3) show the evolution of the solution for $n^0 = 0.3$. The two phases that are the spinodal decomposition and the coarsening are retrieved. Between the Figures 5(b1) and 5(b2), we observe that the solution evolves from a random uniform configuration to an organization in small aggregates that are not saturated. Then (Fig. 5(b3)), the cell density is distributed in elongated and saturated aggregates. The separation of the two phases is clear. However, using Figure 6 (middle), we observe that at the end of the simulation the cell density continues to rearrange. Due to the degeneracy of the mobility, this evolution is very slow.

In Figures 5(c1, c2, c3), we can observe the evolution of the solution for $n^0 = 0.36$. Again, the solution goes through the spinodal decomposition and coarsening phases. The only difference that needs to be highlighted for this simulation is the different shape of the aggregates at the end. Indeed, the initial mass being $n^0 = 0.36 > n^*/2$, the aggregates are wider and more connected to each others.

Therefore, depending on the initial mass of cells in the domain, the 2D simulations of the model show very different spatial organizations of the cell density.

Compared to the reference work [2], the organizations of the cells for the different initial cell densities are the same. No clear difference can be established regarding the simulation involving the relaxed model and the original one.

The three regions corresponding to a steady-state described in [32] are retrieved at the end of the simulations for these 2D test cases.

5.2. Effect of the relaxation parameter σ

In this section we evaluate the effect of the relaxation parameter σ for the stability of the scheme, and in particular to satisfy the CFL-like condition (4.5). This conditions is necessary to preserve the nonnegativity of the solutions of the linear discrete scheme. To evaluate the effect of this parameter on the choice of the time step Δt , we compute the amplification matrix H defined by

$$X^{k+1} = HX^k, \quad \text{with} \quad X^k = \begin{bmatrix} n^k \\ \varphi^k \end{bmatrix}.$$

Here, X^k is called the state vector. Using the matrix form of the scheme (4.1a) and (4.1b) we can decomposed the amplification matrix by $H = H_1^{-1}H_2$ with

$$H_1 = \begin{bmatrix} 0 & \sigma Q + M_l \\ M_l + \Delta t L^k & \Delta t U^k \end{bmatrix}, \quad H_2 = \begin{bmatrix} \gamma Q - (1 - n^*)M_l & \frac{\sigma}{\gamma}(1 - n^*)M_l \\ M_l & 0 \end{bmatrix}.$$

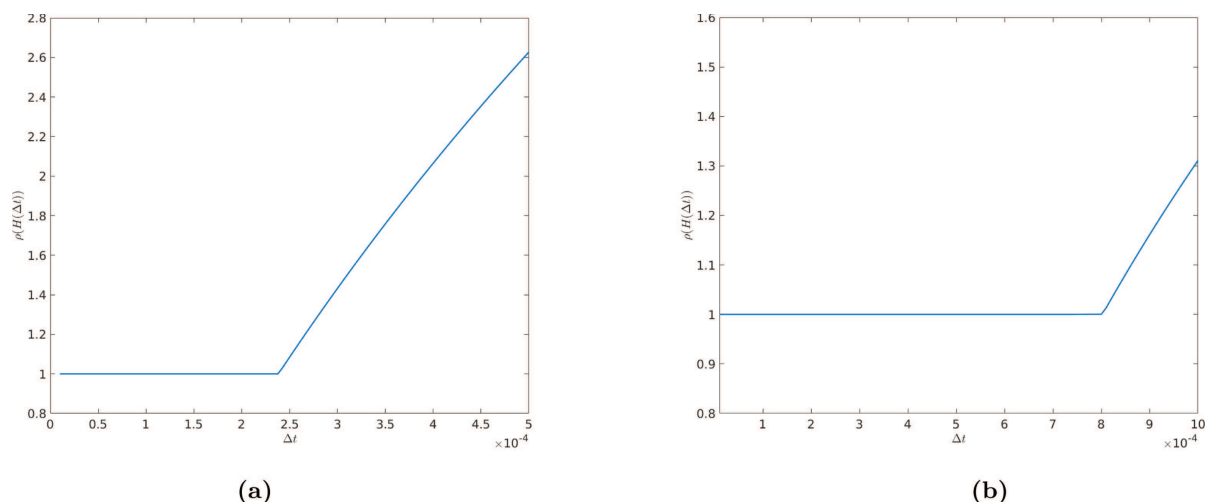


FIGURE 7. Spectral radius as a function of Δt for $\sigma = 10^{-5}$ (left) and $\sigma = 10^{-4}$ (right).

We denote by $\lambda_i, i = 1, \dots, N_h$, the eigenvalues of the amplification matrix H .

To analyze the stability of the numerical scheme due to the relaxation parameter, we compute the spectral radius of the amplification matrix

$$\rho(H(\Delta t)) = \max_i (|\lambda_i|),$$

for a smooth initial conditions. The scheme is stable when the maximum value of the modulus of the eigenvalues is less or equal to 1. Figure 7 represents the spectral radius in function of the time step Δt for two values of σ (the other parameters are the ones taken from the one dimensional test cases with $n^0 = 0.3$). We can observe that the scheme remains stable when Δt is small for the two test cases. However, we see that increasing σ allows to take larger time steps while remaining stable. This result can be explained due to the fact that increasing σ diminishes the value

$$\max_{\substack{x_i \in J_h \\ x_j \in \Lambda_i}} (\varphi_j^{k+1} - \varphi_i^{k+1}),$$

present in the stability condition (4.5). Therefore, the regularization induced by the relaxation parameter allows for faster simulations, but it has an effect on the accuracy of the solution compared to the solution given by the non-relaxed model. However, at the moment it remains unclear how to compare the solution given by a simulation of the relaxed model and a solution of the original degenerate model (without relaxation). This will be the subject of a further work.

6. CONCLUSION

We described and studied a finite element method to solve the relaxed degenerate Cahn–Hilliard equation with single-well logarithmic potential. We considered two time discretization schemes: nonlinear semi-implicit and linear semi-implicit.

We showed that the nonlinear scheme preserves the physical properties of the solutions of the continuous model. We proved that this scheme is well-posed and convergent in dimension $d = 1, 2, 3$. The nonnegativity of the solutions is retrieved thanks to the use of an upwind method adapted within the finite element framework. The linear semi-implicit scheme allows faster simulations and we proved that it is well-posed and preserves the nonnegativity of the solutions as well. We presented some numerical simulations using this linear scheme in one and two dimensions. The numerical simulations validated the nonnegativity-preserving and energy decaying

properties of the scheme. The numerical solutions of the finite element approximation of the RDCH model are in good agreement with previous works dealing with the non-relaxed model. We showed that the relaxation parameter σ allows us to take a larger time step in the scheme (as long as the condition for the nonnegativity is preserved and $\sigma < \gamma$). We point out that thanks to the spatial relaxation, our numerical scheme can be easily implemented and simulations of the relaxed degenerate Cahn–Hilliard model can be computed efficiently using standard softwares.

APPENDIX A. PROOF OF THEOREM 3.14

First, we choose a test function $\chi \in C_c^\infty(\Omega_T, \mathbb{R}^+)$ with $\chi(T, \cdot) = 0$. Then, we multiply equation (3.9) by $\Delta t_m \chi(x_i, t^k) = \Delta t_m \chi_i^k$, sum over the $x_i \in J_h$ and over the $k = 0, \dots, N_T - 1$ to obtain

$$B_m + C_m + D_m + E_m + F_m + G_m = 0,$$

with

$$B_m = \sum_{k=0}^{N_T-1} (n_h^{k+1} - n_h^k, \chi_h^k)^h, \quad (\text{A.1})$$

$$C_m = \frac{\gamma}{\sigma} \sum_{k=0}^{N_T-1} \Delta t_m \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} |Q_{ij}| \left(\hat{B}_{ij}^{k+1} (n_i^{k+1} - n_j^{k+1}) - \sqrt{\hat{B}_{ij}^{k+1}} (\eta_i^{k+1} - \eta_j^{k+1}) \right) (\chi_i^k - \chi_j^k), \quad (\text{A.2})$$

$$D_m = \frac{\gamma}{\sigma} \sum_{k=0}^{N_T-1} \Delta t_m \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} \sqrt{\hat{B}_{ij}^{k+1}} (\eta_i^{k+1} - \eta_j^{k+1}) (\chi_i^k - \chi_j^k), \quad (\text{A.3})$$

$$E_m = \sum_{k=0}^{N_T-1} \Delta t_m \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} |Q_{ij}| \left((\widehat{B\psi''})_{ij}^{k+1} (n_i^{k+1} - n_j^{k+1}) - \sqrt{(\widehat{B\psi''})_{ij}^{k+1}} (\zeta_i^{k+1} - \zeta_j^{k+1}) \right) (\chi_i^k - \chi_j^k), \quad (\text{A.4})$$

$$F_m = \sum_{k=0}^{N_T-1} \Delta t_m \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} \sqrt{(\widehat{B\psi''})_{ij}^{k+1}} (\zeta_i^{k+1} - \zeta_j^{k+1}) (\chi_i^k - \chi_j^k), \quad (\text{A.5})$$

$$G_m = -\frac{\gamma}{\sigma} \sum_{k=0}^{N_T-1} \Delta t_m \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} \tilde{B}_{ij}^{k+1} |Q_{ij}| (w_i^{k+1} - w_j^{k+1}) (\chi_i^k - \chi_j^k). \quad (\text{A.6})$$

Next, we show the limit of each quantity when $m \rightarrow \infty$.

For B_m , since $\chi_h^{N_T} = 0$, we have (see [24])

$$\begin{aligned} B_m &= \sum_{k=0}^{N_T-1} (n_h^{k+1}, \chi_h^k)^h - \sum_{k=1}^{N_T} (n_h^k, \chi_h^k)^h - (n_h^0, \chi_h^0)^h \\ &= - \sum_{k=0}^{N_T-1} \Delta t_m \left(n_h^{k+1}, \frac{\chi_h^{k+1} - \chi_h^k}{\Delta t_m} \right)^h - (n_h^0, \chi_h^0)^h. \end{aligned}$$

From the strong convergence (3.25) and the regularity of χ , we obtain

$$B_m \rightarrow - \int_{\Omega_T} n \frac{\partial \chi}{\partial t} \, dx \, dt - (n^0, \chi^0), \quad \text{as } m \rightarrow \infty.$$

For C_m , we aim to show that $C_m \rightarrow 0$ as $m \rightarrow \infty$. To do so, we rewrite

$$C_m = \frac{\gamma}{\sigma} \sum_{k=0}^{N_T-1} \Delta t_m \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} |Q_{ij}| \sqrt{\hat{B}_{ij}^{k+1}} \left(\sqrt{\hat{B}_{ij}^{k+1}} - \sqrt{\bar{B}_{ij}^{k+1}} \right) (n_i^{k+1} - n_j^{k+1}) (\chi_i^k - \chi_j^k),$$

where

$$\bar{B}_{ij}^{k+1} = \begin{cases} \left(\frac{\eta_i^{k+1} - \eta_j^{k+1}}{n_i^{k+1} - n_j^{k+1}} \right)^2, & \text{if } n_i^{k+1} \neq n_j^{k+1}, \\ \hat{B}_i^{k+1}, & \text{if } n_i^{k+1} = n_j^{k+1}. \end{cases}$$

From the Cauchy–Schwarz Inequality, we have

$$\begin{aligned} |C_m| &\leq \frac{\gamma}{\sigma} \left(\sum_{k=0}^{N_T-1} \Delta t_m \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} |Q_{ij}| \hat{B}_{ij}^{k+1} (n_i^{k+1} - n_j^{k+1})^2 \right)^{1/2} \\ &\quad \times \left(\sum_{k=0}^{N_T-1} \Delta t_m \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} |Q_{ij}| \left(\sqrt{\hat{B}_{ij}^{k+1}} - \sqrt{\bar{B}_{ij}^{k+1}} \right)^2 (\chi_i^{k+1} - \chi_j^{k+1})^2 \right)^{1/2}. \end{aligned}$$

The first term on the right-hand side is bounded from (3.15). The second term, that we denote by R_m , is handled using Lemma A.1 in [14]. Indeed, we denote for each $K \in \mathcal{T}^{h_m}$,

$$\bar{\eta}_K^{k+1} = \max_{x_i \in K} (\eta_h^{k+1}), \quad \underline{\eta}_K^{k+1} = \min_{x_i \in K} (\eta_h^{k+1}),$$

and we define the uniformly continuous function $\sqrt{b \circ \eta}$, defined on the interval $[0, \eta(1)]$, such that for each $K \in \mathcal{T}^{h_m}$,

$$\left| \sqrt{\hat{B}_{ij}^{k+1}} - \sqrt{\bar{B}_{ij}^{k+1}} \right| \leq \omega(\bar{\eta}_K^{k+1} - \underline{\eta}_K^{k+1}), \quad \forall x_j \in K \cap \Lambda_i.$$

Therefore, we have

$$\begin{aligned} R_m &= \sum_{k=0}^{N_T-1} \Delta t_m \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} |Q_{ij}| \left(\sqrt{\hat{B}_{ij}^{k+1}} - \sqrt{\bar{B}_{ij}^{k+1}} \right)^2 (\chi_i^{k+1} - \chi_j^{k+1})^2 \\ &\leq \sum_{k=0}^{N_T-1} \Delta t_m \sum_{K \in \mathcal{T}^{h_m}} \left(\omega(\bar{\eta}_K^{k+1} - \underline{\eta}_K^{k+1}) \right)^2 \sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} |Q_{ij}| (\chi_i^{k+1} - \chi_j^{k+1})^2. \end{aligned}$$

From the regularity of χ , we know that $\sum_{i=0}^{N_h} \sum_{x_j \in \Lambda_i} |Q_{ij}| (\chi_i^{k+1} - \chi_j^{k+1})^2$ is bounded and we obtain

$$R_m \leq C \int_{\Omega_T} \omega(\bar{\eta}_K^{k+1} - \underline{\eta}_K^{k+1}) \, dx \, dt.$$

From Lemma A.1 in [14], we obtain $R_m \rightarrow 0$, as $m \rightarrow \infty$, hence, $C_m \rightarrow 0$. Using similar arguments, we have $E_m \rightarrow 0$, as $m \rightarrow \infty$.

We now turn to the convergence of D_m . To this end, we define two piecewise constant quantities on each element $K \in \mathcal{T}^{h_m}$

$$\Theta_{h_m, \Delta t_m} = \sqrt{b \circ \eta^{-1}}(\Gamma_{h_m, \Delta t_m}), \quad \text{and} \quad \Gamma_{h_m, \Delta t_m}(t, x) = \eta_{h_m, \Delta t_m}(t, x_T), \quad \forall x \in K, \quad t \in (t^k, t^{k+1}].$$

Then, we define

$$\begin{aligned} D'_m &= \frac{\gamma}{\sigma} \int_{\Omega_T} \Theta_{h_m, \Delta t_m} \nabla \eta_{h_m, \Delta t_m} \cdot \nabla \chi_{h_m, \Delta t_m} \, dx \, dt, \\ &= \frac{\gamma}{\sigma} \sum_{k=0}^{N_T-1} \Delta t_m \sum_{K \in \mathcal{T}^{h_m}} \Theta_{h_m, \Delta t_m} \sum_{x_i \in K} \sum_{x_j \in K \cap \Lambda_i} |Q_{ij}| (\eta_i^{k+1} - \eta_j^{k+1}) (\chi_i^k - \chi_j^k). \end{aligned}$$

From a generalization of Lemma A.1 in [14], and the boundedness of the two previous quantities, we have, as $m \rightarrow \infty$,

$$\begin{aligned}\Theta_{h_m, \Delta t_m} &\rightarrow \sqrt{b(n)}, & \text{in } L^2(\Omega_T), \\ \Gamma_{h_m, \Delta t_m} &\rightarrow \eta(n), & \text{in } L^2(\Omega_T),\end{aligned}$$

from which we conclude

$$D'_m \rightarrow \frac{\gamma}{\sigma} \int_{\Omega_T} \sqrt{b(n)} \nabla \eta(n) \cdot \nabla \chi \, dx \, dt, \quad \text{as } m \rightarrow \infty.$$

However, we need to show that $|D_m - D'_m| \rightarrow 0$ as $m \rightarrow \infty$. We use similar arguments as to show that $C_m \rightarrow 0$. Indeed, we rewrite

$$D_m - D'_m = \frac{\gamma}{\sigma} \sum_{k=0}^{N_T-1} \Delta t_m \sum_{i=0}^{N_h} \sum_{K \in \mathcal{T}^{h_m}: x_i \in K} \sum_{x_j \in K \cap \Lambda_i} |Q_{ij}| \left(\sqrt{\hat{B}_{ij}^{k+1}} - \sqrt{a_K^{k+1}} \right) (\eta_i^{k+1} - \eta_j^{k+1})(\chi_i^k - \chi_j^k)$$

with $\sqrt{a_K^{k+1}} = \Theta_K^{k+1}$, we are in position to repeat the arguments presented for the convergence of $C_m \rightarrow 0$ and we do not repeat them here. Therefore, we obtain, as $m \rightarrow \infty$,

$$D_m \rightarrow \frac{\gamma}{\sigma} \int_{\Omega_T} \sqrt{b(n)} \nabla \eta(n) \cdot \nabla \chi \, dx \, dt.$$

The same arguments are also applied to the convergence of F_m to obtain, as $m \rightarrow \infty$,

$$F_m \rightarrow \int_{\Omega_T} \sqrt{b(n) \psi_+''(n)} \nabla \zeta(n) \cdot \nabla \chi \, dx \, dt.$$

The last term of the first equation to analyze is G_m . We repeat again the same arguments. We use the previously defined quantities $\Theta_{h_m, \Delta t_m}$ and $\Gamma_{h_m, \Delta t_m}$. Then, we define

$$\begin{aligned}G'_m &= \frac{\gamma}{\sigma} \int_{\Omega_T} (\Theta_{h_m, \Delta t_m})^2 \nabla w_{h_m, \Delta t_m} \cdot \nabla \chi_{h_m, \Delta t_m} \, dx \, dt, \\ &= \frac{\gamma}{\sigma} \sum_{k=0}^{N_T-1} \Delta t_m \sum_{K \in \mathcal{T}^{h_m}} (\Theta_{h_m, \Delta t_m})^2 \sum_{x_i \in K} \sum_{x_j \in K \cap \Lambda_i} |Q_{ij}| (w_i^{k+1} - w_j^{k+1})(\chi_i^k - \chi_j^k),\end{aligned}$$

and we have, as $m \rightarrow \infty$,

$$G'_m \rightarrow \frac{\gamma}{\sigma} \int_{\Omega_T} b(n) \nabla w \cdot \nabla \chi \, dx \, dt.$$

We still have to prove that $D_m - D'_m \rightarrow 0$ as $m \rightarrow \infty$. Since we have

$$\left| \tilde{B}_{ij}^{k+1} - (\Theta_{h_m, \Delta t_m})^2 \right| \leq C \omega^2 \left| \bar{\eta}_K^{k+1} - \underline{\eta}_K^{k+1} \right|^2,$$

we obtain the result using similar arguments as for the convergence of R_m . In the limit $m \rightarrow \infty$, we obtained the first equation of the weak system defined in (1.1).

For the second equation, we obtain the limit equation using the weak convergence of $w_{h_m, \Delta t_m}$ given by (3.26), the strong convergence of $w_{h_m, \Delta t_m}$ given by (3.25), the continuity of $\psi'_-(\cdot)$, and the strong convergence of $n_{h_m, \Delta t_m}$ given by (3.27).

This concludes the proof of Theorem 3.14.

Acknowledgements. A.P. would like to thank Professor Katharina Schratz for useful discussions and valuable comments. A.P. has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 740623).

REFERENCES

- [1] A. Agosti, Discontinuous Galerkin finite element discretization of a degenerate Cahn–Hilliard equation with a single-well potential. *Calcolo* **56** (2019) 1–47.
- [2] A. Agosti, P.F. Antonietti, P. Ciarletta, M. Grasselli and M. Verani, A Cahn–Hilliard-type equation with application to tumor growth dynamics. *Math. Methods Appl. Sci.* **40** (2017) 7598–7626.
- [3] A. Agosti, S. Marchesi, G. Scita and P. Ciarletta, Modelling cancer cell budding in-vitro as a self-organised, non-equilibrium growth process. *J. Theor. Biol.* **492** (2020) 110203.
- [4] L. Almeida, F. Bubba, B. Perthame and C. Pouchol, Energy and implicit discretization of the Fokker–Planck and Keller–Segel type equations. *Netw. Heterog. Media* **14** (2019) 23–41.
- [5] K. Baba and M. Tabata, On a conservative upwind finite element scheme for convective diffusion equations. *RAIRO Anal. Numér.* **15** (1981) 3–25.
- [6] J.W. Barrett, J.F. Blowey and H. Garcke, Finite element approximation of the Cahn–Hilliard equation with degenerate mobility. *SIAM J. Numer. Anal.* **37** (1999) 286–318.
- [7] M. Bessemoulin-Chatard and A. Jüngel, A finite volume scheme for a Keller–Segel model with additional cross-diffusion. *IMA J. Numer. Anal.* **34** (2014) 96–122.
- [8] S.C. Brenner and L.R. Scott, The Mathematical Theory of Finite Element Methods, 3rd edition. Vol. 15 of *Texts in Applied Mathematics*. Springer, New York (2008).
- [9] S.C. Brenner, S. Gu, T. Gudi and L.-Y. Sung, A quadratic C° interior penalty method for linear fourth order boundary value problems with boundary conditions of the Cahn–Hilliard type. *SIAM J. Numer. Anal.* **50** (2012) 2088–2110.
- [10] F. Bubba, C. Pouchol, N. Ferrand, G. Vidal, L. Almeida, B. Perthame and M. Sabbah, A chemotaxis-based explanation of spheroid formation in 3d cultures of breast cancer cells. *J. Theor. Biol.* **479** (2019) 73–80.
- [11] H. Byrne and L. Preziosi, Modelling solid tumour growth using the theory of mixtures. *Math. Med. Biol.* **20** (2003) 341–366.
- [12] J.W. Cahn, On spinodal decomposition. *Acta Metall.* **9** (1961) 795–801.
- [13] J.W. Cahn and J.E. Hilliard, Free energy of a nonuniform system. I. Interfacial free energy. *J. Chem. Phys.* **28** (1958) 258–267.
- [14] C. Cancès and C. Guichard, Convergence of a nonlinear entropy diminishing control volume finite element scheme for solving anisotropic degenerate parabolic equations. *Math. Comput.* **85** (2016) 549–580.
- [15] C. Cancès, M. Ibrahim and M. Saad, Positive nonlinear CVFE scheme for degenerate anisotropic Keller–Segel system. *SMAI J. Comput. Math.* **3** (2017) 1–28.
- [16] J.A. Carrillo, S. Hittmeir and A. Jüngel, Cross diffusion and nonlinear diffusion preventing blow up in the Keller–Segel model. *Math. Models Methods Appl. Sci.* **22** (2012) 1250041.
- [17] C. Chatelain, T. Balois, P. Ciarletta and M. Ben Amar, Emergence of microstructural patterns inskin cancer: A phase separation analysis in a binary mixture. *New J. Phys.* **13** (2011) 339–357.
- [18] L. Cherfil, A. Miranville and S. Zelik, The Cahn–Hilliard equation with logarithmic potentials. *Milan J. Math.* **79** (2011) 561–596.
- [19] M. Ebenbeck and H. Garcke, On a Cahn–Hilliard–Brinkman model for tumor growth and its singular limits. *SIAM J. Math. Anal.* **51** (2019) 1868–1912.
- [20] C.M. Elliott, The Cahn–Hilliard model for the kinetics of phase separation. In: *Mathematical Models for Phase Change Problems* (Óbidos, 1988). Vol. 88 of *International Series of Numerical Mathematics*. Birkhäuser, Basel (1989) 35–73.
- [21] C.M. Elliott and H. Garcke, On the Cahn–Hilliard equation with degenerate mobility. *SIAM J. Math. Anal.* **27** (1996) 404–423.
- [22] C.M. Elliott and Z. Songmu, On the Cahn–Hilliard equation. *Arch. Ration. Mech. Anal.* **96** (1986) 339–357.
- [23] C.M. Elliott, D.A. French and F.A. Milner, A second order splitting method for the Cahn–Hilliard equation. *Numer. Math.* **54** (1989) 575–590.
- [24] R. Eymard, T. Gallouët, R. Herbin and A. Michel, Convergence of a finite volume scheme for nonlinear degenerate parabolic equations. *Numer. Math.* **92** (2002) 41–82.
- [25] R. Eymard, R. Herbin and A. Michel, Mathematical study of a petroleum-engineering scheme. *ESAIM: M2AN* **37** (2003) 937–972.
- [26] D.J. Eyre, An unconditionally stable one-step scheme for gradient systems (1997).
- [27] H. Fujii, Some remarks on finite element analysis of time dependent field problems. In: *Theory and Practice in Finite Element Structural Analysis*, edited by Y. Yamada and R.H. Gallager. University of Tokyo Press (1973) 91–106.
- [28] H. Garcke, K.F. Lam, E. Sitka and V. Styles, A Cahn–Hilliard–Darcy model for tumour growth with chemotaxis and active transport. *Math. Models Methods Appl. Sci.* **26** (2016) 1095–1148.
- [29] H. Garcke, K.F. Lam, R. Nürnberg and E. Sitka, A multiphase Cahn–Hilliard–Darcy model for tumour growth with necrosis. *Math. Models Methods Appl. Sci.* **28** (2018) 525–577.
- [30] A. Miranville, The Cahn–Hilliard equation and some of its variants. *AIMS Math.* **2** (2017) 479–544.
- [31] A. Novick-Cohen, The Cahn–Hilliard equation. In: *Handbook of Differential Equations: Evolutionary Equations*. Vol. IV of *Handbook of Differential Equations*. Elsevier/North-Holland, Amsterdam (2008) 201–228.
- [32] B. Perthame and A. Poulain, Relaxation of the Cahn–Hilliard equation with singular single-well potential and degenerate mobility. *Eur. J. Appl. Math.* **32** (2021) 89–112.
- [33] A. Quarteroni and A. Valli, Numerical Approximation of Partial Differential Equations. Vol. 23 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin (1994).

- [34] V. Thomée, Galerkin Finite Element Methods for Parabolic Problems. Vol. 25 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin (1997).
- [35] G. Tierra and F. Guillén-González, Numerical methods for solving the Cahn–Hilliard equation and its applicability to related energy-based models. *Arch. Comput. Methods Eng.* **22** (2015) 269–289.
- [36] J.M. Varah, A lower bound for the smallest singular value of a matrix. *Linear Algebra Appl.* **11** (1975) 3–5.
- [37] S.M. Wise, J.S. Lowengrub, H.B. Frieboes and V. Cristini, Three-dimensional multispecies nonlinear tumor growth – I: Model and numerical method. *J. Theor. Biol.* **253** (2008) 524–543.

Subscribe to Open (S2O)

A fair and sustainable open access model



This journal is currently published in open access under a Subscribe-to-Open model (S2O). S2O is a transformative model that aims to move subscription journals to open access. Open access is the free, immediate, online availability of research articles combined with the rights to use these articles fully in the digital environment. We are thankful to our subscribers and sponsors for making it possible to publish this journal in open access, free of charge for authors.

Please help to maintain this journal in open access!

Check that your library subscribes to the journal, or make a personal donation to the S2O programme, by contacting subscribers@edpsciences.org

More information, including a list of sponsors and a financial transparency report, available at: <https://www.edpsciences.org/en/maths-s2o-programme>